



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 144963

TO: Bao-Qun Li  
Location: rem/3d24/3c18  
Art Unit: 1648  
Tuesday, February 22, 2005

Case Serial Number: 09/664363

From: Edward Hart  
Location: Biotech-Chem Library  
REM-1A55  
Phone: 571-272-2512

[edward.hart@uspto.gov](mailto:edward.hart@uspto.gov)

### Search Notes

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**This Page Blank (uspto)**

STIC-Biotech/ChemLib

144963

mej

From: Li, Bao-Qun  
Sent: Friday, February 11, 2005 8:10 AM  
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF SEQ ID NO; 4 IN APPLICATION SN.  
09,664,363. THANKS.

Bao Qun Li M.D  
TC 1600  
Art Unit 1648  
Tel. 517-272-0904  
REM, 3C18  
Rm. 3D24

RECEIVED  
FEB 11 2005  
(STIC)

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: 2/11/05  
Date Completed: 2/22/05  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA Sequence: # 1  
AA Sequence: # \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: STIC  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**





# STIC SEARCH RESULTS

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



**This Page Blank (uspto)**



```
OY 61 GAGAGGTCACTTCCAGGTCGGGCTCAACCAATACCTGATGGGTGGAGCTCCATGC 120
DB 61 GAGAGGTCACTTCCAGGTCGGGCTCAACCAATACCTGATGGGTGGAGCTCCATGC 120
OY 121 GAGCCGGAACCGGATGTAGAGTGTCTCACTTCCATGCTCAACGACCTCCCATGACA 180
DB 121 GAGCCGGAACCGGATGTAGAGTGTCTCACTTCCATGCTCAACGACCTCCCATGACA 180
OY 181 GCAGAGACGGCTAAGGCGAGGCTGGCCAGAGGGGTCTCCCGCTCTTGACAGCTCTTCA 240
DB 181 GCAGAGACGGCTAAGGCGAGGCTGGCCAGAGGGGTCTCCCGCTCTTGACAGCTCTTCA 240
OY 241 GCTAGCAGATTGTCTGGCCCTTCTCTGAAAGCGACATATTAATCAATTAATGCTTCCCA 300
DB 241 GCTAGCAGATTGTCTGGCCCTTCTCTGAAAGCGACATATTAATCAATTAATGCTTCCCA 300
OY 301 GAGCTGACCTCATCGAGGCTCAACCTCTGTGTGGCGGACATGAGTGGCGGGGACATTAC 360
DB 301 GAGCTGACCTCATCGAGGCTCAACCTCTGTGTGGCGGACATGAGTGGCGGGGACATTAC 360
OY 361 CGCGTGAAGTCAGAGAACAGAGTATCTGGAATCTTTTGACCCGCTCCGAGCGAG 420
DB 361 CGCGTGAAGTCAGAGAACAGAGTATCTGGAATCTTTTGACCCGCTCCGAGCGAG 420
OY 421 GAGGATGACCGGGAAGTGTCCGTCCTCGGCGAGATCTCGGAAATCCAAAGAAATTCCTCA 480
DB 421 GAGGATGACCGGGAAGTGTCCGTCCTCGGCGAGATCTCGGAAATCCAAAGAAATTCCTCA 480
OY 481 CCAAGCATGCTCCGATGCGGACGCGCGGATTAACACCTTCGCTGTGAGATCTTGAAG 540
DB 481 CCAAGCATGCTCCGATGCGGACGCGCGGATTAACACCTTCGCTGTGAGATCTTGAAG 540
OY 541 GCCCGGACCTACGTCCTCCAGTGTGACATGAGTGGTCCCATCTGCACTTAAGACCTT 600
DB 541 GCCCGGACCTACGTCCTCCAGTGTGACATGAGTGGTCCCATCTGCACTTAAGACCTT 600
OY 601 CCTATTAACAATCTCAAGGAAAGAGACAGTGTCTGACAGAAATCCACCTGTCTTCT 660
DB 601 CCTATTAACAATCTCAAGGAAAGAGACAGTGTCTGACAGAAATCCACCTGTCTTCT 660
OY 661 GCGCTGCGGAGCTTCCCAAAAGGCTTTTGGTGTAGCTCGGACCGTCCGCTCGACAGC 720
DB 661 GCGCTGCGGAGCTTCCCAAAAGGCTTTTGGTGTAGCTCGGACCGTCCGCTCGACAGC 720
OY 721 GGCACGCGCAACGCGCTCTGACCAATCTCTCGACGAGCGGAGGACAGATCTGACGTT 780
DB 721 GGCACGCGCAACGCGCTCTGACCAATCTCTCGACGAGCGGAGGACAGATCTGACGTT 780
OY 781 GAGTGTATCTCTCAATGCCCCCTTTGAGGGGAGCGCGGAGACCCCGATCTCAGCGAC 840
DB 781 GAGTGTATCTCTCAATGCCCCCTTTGAGGGGAGCGCGGAGACCCCGATCTCAGCGAC 840
OY 841 GGGTCTTGTGTACCGTGTAGAGGAGCGGCTGTAGAGAGTGTGTGTGTGTGTGTGTTC 900
DB 841 GGGTCTTGTGTACCGTGTAGAGGAGCGGCTGTAGAGAGTGTGTGTGTGTGTGTGTTC 900
OY 901 TACACATGAGACGCGCTCTGATCAAGCATGCGCTCGGAGGAAAGCAAGTGTGCCATC 960
DB 901 TACACATGAGACGCGCTCTGATCAAGCATGCGCTCGGAGGAAAGCAAGTGTGCCATC 960
OY 961 AACGCGTTGAGCAACTTTGTGTGCTCAACAACATGATCTACGCTACCAATCCGCG 1020
DB 961 AACGCGTTGAGCAACTTTGTGTGCTCAACAACATGATCTACGCTACCAATCCGCG 1020
OY 1021 AGCGCAAGCCAGCG 1035
DB 1021 AGCGCAAGCCAGCG 1035
```

RESULT 2  
LOCUS A32185 1035 bp RNA linear PAT 10-DEC-1996

```
DEFINITION NANBHV PT polyprotein fragment.
ACCESSION A32185
VERSION A32185.1 GI:1926540
KEYWORDS
SOURCE Non-A, non-B hepatitis virus
ORGANISM Non-A, non-B hepatitis virus
REFERENCE Viruses; unclassified viruses.
AUTHORS 1 (bases 1 to 1035)
JOURNAL Patent: FR 2655990-A 4 21-JUN-1991;
FEATURES
source
Location/Qualifiers
1..1035
/organism="Non-A, non-B hepatitis virus"
/mol_type="unassigned RNA"
/db_xref="taxon:12440"
/clone="JG3"
/clone_1db="lambda gt11"
<1..>1035
/codon_start=1
/product="partial NANBH-PT polyprotein"
/protein_id="CA02089.1"
/db_xref="GI:1926541"
/translation="TEVDGVLRHRYAACKPLREBYTFOYGLNOYLVSOLPREPER
DVAVLTSMLDPSHTATARKRLAKGSPESLASSASQUSGPKATYITQNPDA
DLIBANLIRHEMGGDITRVESENKVLIDFPDLRAEDREVSPAEILRSKEP
PAMPAMARPDYNPPLBSWKADVPVPHVHCPLPTKTPPIPPRRKRTVLSTEV
SSALAEIATKAFGSSGSAVDSGTATAPDQSDGGSVSSMPPLEGKPD
DLSGWSVTEBEAGBDVCCSMGYTMTGALITPCAEBSLPIINALSNSLRHNMV
YATTSRSASQR"
ORIGIN
Query Match 100.0%; Score 1035; DB 6; Length 1035;
Best Local Similarity 100.0%; Pred. No. 1.5e-228;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ACAGAGTGAATGGGGTCCGCTGCACAGTACGCTCCGCGTGCAAACCTTCTCTACGG 60
DB 1 ACAGAGTGAATGGGGTCCGCTGCACAGTACGCTCCGCGTGCAAACCTTCTCTACGG 60
OY 61 GAGAGGTCACTTCCAGGTCGGGCTCAACCAATACCTGATGGGTGGAGCTCCATGC 120
DB 61 GAGAGGTCACTTCCAGGTCGGGCTCAACCAATACCTGATGGGTGGAGCTCCATGC 120
OY 121 GAGCCGGAACCGGATGTAGAGTGTCTCACTTCCATGCTCAACGACCTCCCATGACA 180
DB 121 GAGCCGGAACCGGATGTAGAGTGTCTCACTTCCATGCTCAACGACCTCCCATGACA 180
OY 181 GCAGAGACGGCTAAGCGGAGCTGGCCAGGGGCTCTCCCGCTCTTGACAGCTCTTCA 240
DB 181 GCAGAGACGGCTAAGCGGAGCTGGCCAGGGGCTCTCCCGCTCTTGACAGCTCTTCA 240
OY 241 GCTAGCAGATTGTCTGGCCCTTCTCTGAAAGCGACATATTAATCAATTAATGCTTCCCA 300
DB 241 GCTAGCAGATTGTCTGGCCCTTCTCTGAAAGCGACATATTAATCAATTAATGCTTCCCA 300
OY 301 GACGCTGACCTCATCGAGGCTCAACCTCTGTGTGGCGGACATGAGTGGCGGGGACATTAC 360
DB 301 GACGCTGACCTCATCGAGGCTCAACCTCTGTGTGGCGGACATGAGTGGCGGGGACATTAC 360
OY 361 CGCGTGAAGTCAGAGAACAGTATCTCTGGAATCTTTTGAACCGGCTCCGAGCGAG 420
DB 361 CGCGTGAAGTCAGAGAACAGTATCTCTGGAATCTTTTGAACCGGCTCCGAGCGAG 420
OY 421 GAGGATGACGGGGAAGTGTCCGTCCTCGGCGGAGATCTCGGGAATCCAAAGAAATTCCTCA 480
DB 421 GAGGATGACGGGGAAGTGTCCGTCCTCGGCGGAGATCTCGGGAATCCAAAGAAATTCCTCA 480
OY 481 CCAAGCATGCTCCGATGCGGACGCGCGGATTAACACCTTCGCTGTGAGTCTCTGAGAG 540
DB 481 CCAAGCATGCTCCGATGCGGACGCGCGGATTAACACCTTCGCTGTGAGTCTCTGAGAG 540
OY 541 GCCCGGACCTACGTCCTCCAGTGTGACATGAGTGGTCCCATCTGCACTTAAGACCTT 600
DB 541 GCCCGGACCTACGTCCTCCAGTGTGACATGAGTGGTCCCATCTGCACTTAAGACCTT 600
```

Db 541 GCCCGGAGTACGTCCTCCAGTGTATCATGGGTGCCCACTTAAGACCCCT 600  
QY CTTATACACTCTCAAGGAGAAAGAGACGTTGTTTGACAGATCAACCGTCTTCT 660  
Db 601 CTTATACACTCTCAAGGAGAAAGAGACGTTGTTTGACAGATCAACCGTCTTCT 660  
QY GCCCTGAGGAGCTTGCCCAAAAGGCTTTTGTAGTCCGAGACCGTGGCGTGCAGC 720  
Db 661 GCCCTGAGGAGCTTGCCCAAAAGGCTTTTGTAGTCCGAGACCGTGGCGTGCAGC 720  
QY 721 GGCAGCGCAACCGCCCTCTCTGACCAATCTCCAGCAGCGAGAGAGATCTGAGT 780  
Db 721 GGCAGCGCAACCGCCCTCTCTGACCAATCTCCAGCAGCGAGAGATCTGAGT 780  
QY 781 GAGTCGTATCTCTCAATGCCCCCTTGAGGGGAGCGCGGGAGACCCGATCTCAGGAC 840  
Db 781 GAGTCGTATCTCTCAATGCCCCCTTGAGGGGAGCGCGGGAGACCCGATCTCAGGAC 840  
QY 841 GGGCTTGGTCTACCGGTGAGTGAAGAGCGGTGAGGACGTCTGCTGCTGATGTC 900  
Db 841 GGGCTTGGTCTACCGGTGAGTGAAGAGCGGTGAGGACGTCTGCTGCTGATGTC 900  
QY 901 TACACATGACAGCGCTCTGATCAGCCATGCGCTGCGAGGAAAGCAAGCTGCCATC 960  
Db 901 TACACATGACAGCGCTCTGATCAGCCATGCGCTGCGAGGAAAGCAAGCTGCCATC 960  
QY 961 AATCGCTTGAACAATCTTTGCTGCTGCTCAGCAACAATGCTTACGCTTACCATCCGC 1020  
Db 961 AATCGCTTGAACAATCTTTGCTGCTGCTCAGCAACAATGCTTACGCTTACCATCCGC 1020  
QY 1021 AGGCAAGCCAGCGG 1035  
Db 1021 AGGCAAGCCAGCGG 1035

RESULT 3  
AR144033  
LOCUS AR144033 1035 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 4 from patent US 6210675.  
ACCESSION AR144033  
VERSION AR144033.1 GI:15105900  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1035)  
AUTHORS Highfield, P. Edmund., Rodgers, B. Colin., Tedder, R. Seton. and  
TITLE PT-NANB hepatitis polypeptides  
JOURNAL Patent: US 6210675 A 4 03-APR-2001;  
FEATURES  
Location/Qualifiers  
1..1035  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 100.0%; Score 1035; DB 6; Length 1035;  
Best Local Similarity 100.0%; Pred. No. 1.5e-228;  
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAGTGTATGGGGTGGCGCTGCAAGGTAGCTCCGCGTGCAGAACTCTCTTAAG 60  
Db 1 ACAGAGTGTATGGGGTGGCGCTGCAAGGTAGCTCCGCGTGCAGAACTCTCTTAAG 60  
QY 61 GAGAGAGTCAATTCAGAGTCCGAGCTCAACCAATACCTGTTGGTGCAGCTCCCATG 120  
Db 61 GAGAGAGTCAATTCAGAGTCCGAGCTCAACCAATACCTGTTGGTGCAGCTCCCATG 120  
QY 121 GAGCCCGAAGCCGATGTAGAGTGTCTCATCTTCCATGCTCACCGACCCCTCCACATCA 180  
Db 121 GAGCCCGAAGCCGATGTAGAGTGTCTCATCTTCCATGCTCACCGACCCCTCCACATCA 180  
QY 181 GCAGAGCGGCTAAGGCGAAGCTGGCCAGGGGGTCTCCCGCTCTTGCGCACTCTTCA 240

Db 181 GCAGAGCGGCTAAGGCGAAGCTGGCCAGGGGGTCTCCCGCTCTTGCGCACTCTTCA 240  
QY 241 GCTAGCCAGTGTCTGAGCCCTTCTCGAAAGCGACATTAATCCAAATATGACTTCCCA 300  
Db 241 GCTAGCCAGTGTCTGAGCCCTTCTCGAAAGCGACATTAATCCAAATATGACTTCCCA 300  
QY 301 GAGCGTGAATCTCAATGAGGCGAACCCTCTGTGGCGGATGAGATGGGGCGGGACATTACC 360  
Db 301 GAGCGTGAATCTCAATGAGGCGAACCCTCTGTGGCGGATGAGATGGGGCGGGACATTACC 360  
QY 361 CGCGTGAAGTCAAGAGAACAAAGTATTAATCTTGACCTTTTCGACCCGCTCGAGCGAG 420  
Db 361 CGCGTGAAGTCAAGAGAACAAAGTATTAATCTTGACCTTTTCGACCCGCTCGAGCGAG 420  
QY 421 GAGATGAGCGGGAGTGTCCGTCCGCGGAGATCTTCGGAATTCAGAAATTTCCCA 480  
Db 421 GAGATGAGCGGGAGTGTCCGTCCGCGGAGATCTTCGGAATTCAGAAATTTCCCA 480  
QY 481 CCAGCGATGCCGATGAGGCGACGCCCGGATTAACAACCTCCGCTGCTGAGTCTGGAG 540  
Db 481 CCAGCGATGCCGATGAGGCGACGCCCGGATTAACAACCTCCGCTGCTGAGTCTGGAG 540  
QY 541 GCCCGGAGTACGTCCCTCAGTGTACATGGGTGCCACTGCGCACTTAAGACCCCT 600  
Db 541 GCCCGGAGTACGTCCCTCAGTGTACATGGGTGCCACTGCGCACTTAAGACCCCT 600  
QY 601 CTTATACACTCTCAAGGAGAAAGAGACGTTGTTTGACAGATCAACCGTCTTCT 660  
Db 601 CTTATACACTCTCAAGGAGAAAGAGACGTTGTTTGACAGATCAACCGTCTTCT 660  
QY 661 GCCCTGAGGAGCTTGCCCAAAAGGCTTTTGTAGTCCGAGACCGTGGCGTGCAGC 720  
Db 661 GCCCTGAGGAGCTTGCCCAAAAGGCTTTTGTAGTCCGAGACCGTGGCGTGCAGC 720  
QY 721 GGCAGCGCAACCGCCCTCTCTGACCAATCTCCAGCAGCGAGAGAGATCTGAGT 780  
Db 721 GGCAGCGCAACCGCCCTCTCTGACCAATCTCCAGCAGCGAGAGAGATCTGAGT 780  
QY 781 GAGTCGTATCTCTCAATGCCCCCTTGAGGGGAGACCCGATCTCAGGAC 840  
Db 781 GAGTCGTATCTCTCAATGCCCCCTTGAGGGGAGACCCGATCTCAGGAC 840  
QY 841 GGGCTTGGTCTACCGGTGAGTGAAGAGCGGTGAGGACGTCTGCTGCTGATGTC 900  
Db 841 GGGCTTGGTCTACCGGTGAGTGAAGAGCGGTGAGGACGTCTGCTGCTGATGTC 900  
QY 901 TACACATGACAGCGCTCTGATCAGCCATGCGCTGCGAGGAAAGCAAGCTGCCATC 960  
Db 901 TACACATGACAGCGCTCTGATCAGCCATGCGCTGCGAGGAAAGCAAGCTGCCATC 960  
QY 961 AATCGCTTGAACAATCTTTGCTGCTGCTCAGCAACAATGCTTACGCTTACCATCCGC 1020  
Db 961 AATCGCTTGAACAATCTTTGCTGCTGCTCAGCAACAATGCTTACGCTTACCATCCGC 1020  
QY 1021 AGGCAAGCCAGCGG 1035  
Db 1021 AGGCAAGCCAGCGG 1035

RESULT 4  
A76573  
LOCUS A76573 1788 bp DNA linear PAT 19-OCT-1999  
DEFINITION Sequence 5 from Patent WO9317110.  
ACCESSION A76573  
VERSION A76573.1 GI:608484  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1788)  
AUTHORS Rodgers, B. C. and Parker, D.  
TITLE A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE

JOURNAL Patent: WO 9317110-A 5 02-SEP-1993;  
WELLCOME FOUND (GB)  
FEATURES Location/Qualifiers  
source 1..1788  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
<1..>1788  
/note="Encodes putative PT-NANBH NS5 protein"  
/codon\_start=1  
/product="PT-NANBH coding sequence"  
/protein\_id="CAB58623.1"  
/db\_xref="GI:6088485"

CDS

1 ACAAGAGTGGAGTGGGCTGCGCTGCACAGGTACGCTCCGCGCTGCAACCTCTCTACCG 60  
1 ACAGAGTGAATGGGGTGGCGCTGCACAGGTACGCTCCGCGCTGCAACCTCTCTACCG 60  
61 GAGGAGGTACACCTCCAGGTCCGGCTCAACCAATACCTGGTGGGTGGAGTCCCATAC 120  
61 GAGGAGGTACACCTCCAGGTCCGGCTCAACCAATACCTGGTGGGTGGAGTCCCATAC 120  
121 GAGCCGGAACCGGATGTACAGTGTCACTTCCATGCTCAGACGACCCCTCCACATCACA 180  
121 GAGCCGGAACCGGATGTACAGTGTCACTTCCATGCTCAGACGACCCCTCCACATCACA 180  
181 GAGAGACCGGCTAAGCGCAGGCTGCGCAGGGGGTCTCCCTCTCTGGCCAGCTTTCA 240  
181 GAGAGACCGGCTAAGCGCAGGCTGCGCAGGGGGTCTCCCTCTCTGGCCAGCTTTCA 240  
241 GCTAGCACTTGTCTGGCCCTTCTCTGAGAGCGACATTAATACCAAAATGACTTCCCA 300  
241 GCTAGCACTTGTCTGGCCCTTCTCTGAGAGCGACATTAATACCAAAATGACTTCCCA 300  
301 GAGGCTGACCTGATCGAGGCGCAACCTCTGTGGCGGATGAGATGGCGGGGACATTACC 360  
301 GAGGCTGACCTGATCGAGGCGCAACCTCTGTGGCGGATGAGATGGCGGGGACATTACC 360  
361 GCGGTGAGTCAAGAGAAAGGTAGTAATCTTGACTCTTTTGACCCGCTCCGAGCGAG 420  
361 GCGGTGAGTCAAGAGAAAGGTAGTAATCTTGACTCTTTTGACCCGCTCCGAGCGAG 420  
421 GAGGATGAGCGGGAAGTGTCCGTCGGGGGAGATCTTCCGGAATTCAGAAATTTCCCA 480  
421 GAGGATGAGCGGGAAGTGTCCGTCGGGGGAGATCTTCCGGAATTCAGAAATTTCCCA 480  
481 CAGGAGTCCCGCATGGGCGACCGCCGGATTACACCTCTCCGCTGAGATCTTGGAAG 540  
481 CAGGAGTCCCGCATGGGCGACCGCCGGATTACACCTCTCCGCTGAGATCTTGGAAG 540  
541 GCGCCGAGTACGTCCTTCAGTGTGATACAGGATGCGCACTGCGCACTTAAGACCCCT 600  
541 GCGCCGAGTACGTCCTTCAGTGTGATACAGGATGCGCACTGCGCACTTAAGACCCCT 600  
601 CCTATACCACTCCACGAGAAAGAGACAGTTGTTCTGACAGAAATCCACCGTGTCTTCT 660  
601 CCTATACCACTCCACGAGAAAGAGACAGTTGTTCTGACAGAAATCCACCGTGTCTTCT 660

ORIGIN

Query Match 100.0%; Score 1035; DB 6; Length 1788;  
Best Local Similarity 100.0%; Pred. No. 1.4e-228;  
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

661 GCCCTGCGGAGCTTGCACAAAGGCTTTGTAGCTCCGACCGTCCGCCGTGCACAGC 720  
661 GCCCTGCGGAGCTTGCACAAAGGCTTTGTAGCTCCGACCGTCCGCCGTGCACAGC 720  
721 GGCACGCAACCGCCCTCTCTGACCAATCTCTCCGACGCGCGGAGCAGATCTGACGTT 780  
721 GGCACGCAACCGCCCTCTCTGACCAATCTCTCCGACGCGCGGAGCAGATCTGACGTT 780  
781 GAGTGTATCTTCATGCCCCCTTGAAGGGGAGCGCGGGGAGCCCGGATCTGACGAC 840  
781 GAGTGTATCTTCATGCCCCCTTGAAGGGGAGCGCGGGGAGCCCGGATCTGACGAC 840  
841 GGGTCTGTAGTACCGGTAGTGAAGAGCGCGGTGAGAGCGTCTGCTGCTGATGTC 900  
841 GGGTCTGTAGTACCGGTAGTGAAGAGCGCGGTGAGAGCGTCTGCTGCTGATGTC 900  
901 TACACATGACAGGCGCTCTGATCAAGCGCATGCGCTGCGAGAAAGCAAGCTGCCATC 960  
901 TACACATGACAGGCGCTCTGATCAAGCGCATGCGCTGCGAGAAAGCAAGCTGCCATC 960  
961 AACGCTGAGCAACTCTTGTGCTGCTGACCAACATGCTCTGCTGCTGATCCGCT 1020  
961 AACGCTGAGCAACTCTTGTGCTGCTGACCAACATGCTCTGCTGCTGATCCGCT 1020  
1021 AGCGCAACCGACCGG 1035  
1021 AGCGCAACCGACCGG 1035  
1021 AGCGCAACCGACCGG 1035

RESULT 5  
A76569  
LOCUS A76569 2790 bp DNA linear PAT 19-OCT-1999  
DEFINITION Sequence 1 from Patent WO9317110.  
ACCESSION A76569  
VERSION A76569.1 GI:6088480  
KEYWORDS  
SOURCE  
ORGANISM  
unidentified  
unclassified.  
REFERENCE 1 (bases 1 to 2790)  
AUTHORS Rodgers,B.C. and Parker,D.  
TITLE A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE  
JOURNAL Patent: WO 9317110-A 1 02-SEP-1993;  
WELLCOME FOUND (GB)  
FEATURES Location/Qualifiers  
source 1..2790  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
1..>2790  
/note="unamed protein product; Fusion of non-structural  
and structural protein sequences of PT-NANBH"

CDS

1..>2790  
/note="unamed protein product; Fusion of non-structural  
and structural protein sequences of PT-NANBH"  
/codon\_start=1  
/protein\_id="CAB58621.1"  
/db\_xref="GI:6088481"  
/translation="MPOYSYRPTTGPDPPLSAETREYDGLRHYAPACKPLREBY  
TFQYGLNOLYVGSQLPCEBPDPVALVLSMLTDPSEIATKRLARGLSPSLASSA  
SOLGSPSKATYITQNDPPADLIENALMRHEMGDITRVESNKVYILDSPPLA  
EEDREVSVPALILKSKFPAPAPADYVNPPLLSMADYVPPVHGCP  
KTPPIPPRRKRTVLTSTVSSALALATAFSGSPSADSTAPADPQSSDDG  
AGSDVESYSSMPLEBGPDPDLSDGSGTYSBAGDVVCCSMTWTGALITPCA  
EESLPIINALSNSILRHHNVYATTSASOROKVTFDLQILDHYODVLKEMKAK  
ASTYKAKLISVEBAKILTPHSASKSGYGAQDVRLNSKAIINIRVEMDLDET  
PIDTTMAKNEVFCVQEBGRKARALIVPDLGVRCEKALIDVYSTLPOAWGSS  
YGFQYSPQAVEFLVNAWKSRTKPMGFAVDTRCDSTVENDIRVESIYCCDLAP  
ARQAIRSLTERLYIGPLTNSKGNCRRCRAGSVLTSGNTLCTYKASAVKKK  
KKTARNTNLRPQDVRPPGGQIVGVVLLPRGRLGVRARTRKTSOPGRQPI  
KARQPEGAMAOPEYPMPLVGNEGMAGWILSPRGSPPMPTDPRRSNLLGVID  
TLTGCPADLMGYIPLVGAFLGAGARALAHGVRVLEDPVNTAGNLPGCSFSLI  
SCLTIPASAYEVRVNSGIIHYTNDCSNSIYETADIMHTPGVPCYRBNSSRCV  
ALPTPLAKDASITPATIRRHVDLLVGAAPCASMYGDLGSAVPEPFLSAGRYGSF  
PGRQEBRTHSLQNP"

## ORIGIN

Query Match 100.0%; Score 1035; DB 6; Length 2790;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-228;  
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ACAGAGTGGATGGGGTGGGCTGACAGGTAAGCTCCGCGGTGCAAACTCTCTCAACG 60
DB 64 ACAGAGTGGATGGGGTGGGCTGACAGGTAAGCTCCGCGGTGCAAACTCTCTCAACG 123
QY 61 GAGAGGTCACTATTCAGGTGCGGCTCAACCAATACCTGGTTGGGTGCGAGCTCCATGC 120
DB 124 GAGAGGTCACTATTCAGGTGCGGCTCAACCAATACCTGGTTGGGTGCGAGCTCCATGC 183
QY 121 GAGCCCGAACCGGATGTAGAGTGTCTCATTTCCATGCTCAACCGACCTCCCATCA 180
DB 184 GAGCCCGAACCGGATGTAGAGTGTCTCATTTCCATGCTCAACCGACCTCCCATCA 243
QY 181 GCAGAGACGGCTAAGGCGAGGCTGGCCAGGGGGTCTCCCTCCCTGGGCACTCTTCA 240
DB 244 GCAGAGACGGCTAAGGCGAGGCTGGCCAGGGGGTCTCCCTCCCTGGGCACTCTTCA 303
QY 241 GCTAGCAGTTGTCTGGCCCTTCTCGAAGCGCATATTAACCAAAATGACTTCCCA 300
DB 304 GCTAGCAGTTGTCTGGCCCTTCTCGAAGCGCATATTAACCAAAATGACTTCCCA 363
QY 301 GAGCTGACCTCATTCGAGGCCAACTTCCTGTGGCGGATGAGTGGCGGGGACATTACC 360
DB 364 GAGCTGACCTCATTCGAGGCCAACTTCCTGTGGCGGATGAGTGGCGGGGACATTACC 423
QY 361 CGGCTGAGTCAAGAAACAAAGTATTAATCTGACCTCTTCAACCGGCTCCGAGCGGAG 420
DB 424 CGGCTGAGTCAAGAAACAAAGTATTAATCTGACCTCTTCAACCGGCTCCGAGCGGAG 483
QY 421 GAGATGAGCGGGAGAGTGTCCGCTCCGCGCGAGATCTCGCGAAATCCAAAGAAATCCCA 480
DB 484 GAGATGAGCGGGAGAGTGTCCGCTCCGCGCGAGATCTCGCGAAATCCAAAGAAATCCCA 543
QY 481 CCAGCGATGCCCGCATGCGGACGCGCGGATTACACCTTCCTGCTGAGTCTCTGGAAG 540
DB 544 CCAGCGATGCCCGCATGCGGACGCGCGGATTACACCTTCCTGCTGAGTCTCTGGAAG 603
QY 541 GCGCCGGAAGTACGCTCCCTCAAGTGTATCATGGGGGCCCATGCGCACTTAAAGCCCT 600
DB 604 GCGCCGGAAGTACGCTCCCTCAAGTGTATCATGGGGGCCCATGCGCACTTAAAGCCCT 663
QY 601 CCTATACCACTCCAGCGAAGAAAGAGAGAGTGTCTGACAGATCCACGCTCTTCT 660
DB 664 CCTATACCACTCCAGCGAAGAAAGAGAGAGTGTCTGACAGATCCACGCTCTTCT 723
QY 661 GCGCTGCGGAGCTTGCACAAAGGCTTTTGTAGCTCCGAGCCGTCGCGCTGACAGC 720
DB 724 GCGCTGCGGAGCTTGCACAAAGGCTTTTGTAGCTCCGAGCCGTCGCGCTGACAGC 783
QY 721 GCGAGCGCAACCGCCCTCTGACCAATCTCCGACGAGCGCGGAGAGATCTGACGTT 780
DB 784 GCGAGCGCAACCGCCCTCTGACCAATCTCCGACGAGCGCGGAGAGATCTGACGTT 843
QY 781 GAGTCGTATTCTCTCATGTCGCGCTTGAAGGGGAGCGCGGAGACCCGATCTCAGCGAC 840
DB 844 GAGTCGTATTCTCTCATGTCGCGCTTGAAGGGGAGCGCGGAGACCCGATCTCAGCGAC 903
QY 841 GGGTCTTGGTCTTACCGTGAAGAGAGCGGTGAGAGAGTCTGTGCTGCTGATGTC 900
DB 904 GGGTCTTGGTCTTACCGTGAAGAGAGCGGTGAGAGAGTCTGTGCTGCTGATGTC 963
QY 901 TACACATGAGCAGGCGCTGATCAAGCGCAATGCGTGGGAGGAGAAAGCAAGTCCCATC 960
DB 964 TACACATGAGCAGGCGCTGATCAAGCGCAATGCGTGGGAGGAGAAAGCAAGTCCCATC 1023
QY 961 AACGCGTTGAGCAACTTGTGCTGCGTCAACCAAACTGATCTTACGCTACCAATCCCGC 1020
DB 1024 AACGCGTTGAGCAACTTGTGCTGCGTCAACCAAACTGATCTTACGCTACCAATCCCGC 1083

```

QY 1021 AGCGCAAGCCAGCGG 1035  
 DB 1084 AGCGCAAGCCAGCGG 1098

## RESULT 6

LOCUS A76594 3372 bp DNA linear PAT 19-OCT-1999  
 DEFINITION Sequence 26 from Patent WO9317110.  
 ACCESSION A76594  
 VERSION A76594.1 GI:6088505

SOURCE unclassified  
 ORGANISM unclassified

REFERENCE 1 (bases 1 to 3372)  
 AUTHORS Rodgers, B.C. and Parker, D.  
 TITLE A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE  
 JOURNAL Patent: WO 9317110-A 26 02-SEP-1993;  
 WELLCOME FOUND (GB)

FEATURES  
 source location/Qualifiers  
 1..3372

CDS  
 1..3372  
 /note="unlabeled protein product; Encodes a fusion protein  
 which can be expressed in insect cells infected with the  
 baculovirus BHC-28"

/protein\_id="CA58632.1"  
 /db\_xref="GI:6088506"

/translation="MPDYSYRPTIGDPDPSLSAEFTVDGVLHRYAPACKPLIREV  
 TFOVGNQVLVGSQLPCEPEPDVAVLVLTMSLTTPSHITABAKRLRSGSPSLASSA  
 SOLGSSSKATYITONDPDADLIBANLIMRHMGGDITRVESENKVLIDSPDRA  
 EDEREVSVPATILRKSKEPPMPAMAPDNPPLIESMKADVPVPHVGGELPPT  
 KHPDIPPRKRTVLTSTGSSALAEATKRGSGGAGVCCMSTTWTGALTTPCA  
 AGSDVBSIMPLEGEPPDSDSNGSTVSEAEEDVCCMSTTWTGALTTPCA  
 ESKLPIINALSNLHNHNVYATTSRSASQOKKQVRLQLDHYDVLKEMAK  
 ASIVKALISVEACKLTTPHSAKFEYGADVRLSKAINHISVWEDLIEDET  
 PIDTTIMAKNEVQPERGGRKPARLIVFDLGVCEKMAIYDVVSTLPQAMGSS  
 YGFOVSGPOVFLVNAKSKTTPMGFAVDTCGFDSTENDIRVRESITYCCDLAP  
 ARALISLTERLIVTIGTFLNSKQNGYRCASGVLCTSCGNTLTCTKASGDIIC  
 DECHSTDSITIGITVLQALTAGRLVLAATPPGAVTPHPNIEEALANTEI  
 PFYKAIPIETIKGGRHLIFCHSKKKCDLAKLVGLINAAVYNGLVSVIPASGD  
 VVVVATDALMTGFTGDPVIDCIVLTQTVFSDPTPTIETTIVPODVAHRSQRG  
 RTGGRGRIYRFVAAYKKKKKTKRNTNLRPODVRPGGQIYGVYLLPRRGRLGV  
 RARTKTSERSOPRGROPIPKAROPGRAMADPGVPMPIYGNMGMACTILSPGRS  
 PSMGPTDPRRSRNLRKVIDTLTTCGADLMGYIPLVGAFLGAAARLARGVRLDGV  
 NTAQNLPCGSPSIFILALLSCUTTPASAYEVKNVSGIYHVNDCNSSTVETADMT  
 MTPGCVPCVRGNSRCWVALPTLAADASIPTATIRRHVDLIVGAAPFCSAMTVG  
 DLGGSVFPEFOLSAGRYGSFPGRORPKTHSLQGNP"

## ORIGIN

Query Match 100.0%; Score 1035; DB 6; Length 3372;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-228;  
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ACAGAGTGGATGGGGTGGGCTGACAGGTAAGCTCCGCGGTGCAAACTCTCTCAACG 60
DB 64 ACAGAGTGGATGGGGTGGGCTGACAGGTAAGCTCCGCGGTGCAAACTCTCTCAACG 123
QY 61 GAGAGGTCACTATTCAGGTGCGGCTCAACCAATACCTGGTTGGGTGCGAGCTCCATGC 120
DB 124 GAGAGGTCACTATTCAGGTGCGGCTCAACCAATACCTGGTTGGGTGCGAGCTCCATGC 183
QY 121 GAGCCCGAACCGGATGTAGAGTGTCTCATTTCCATGCTCAACCGACCTCCCATCA 180
DB 184 GAGCCCGAACCGGATGTAGAGTGTCTCATTTCCATGCTCAACCGACCTCCCATCA 243
QY 181 GCAGAGACGGCTAAGGCGAGGCTGGCCAGGGGGTCTCCCTCCCTGGGCACTCTTCA 240

```

Db 244 GAGAGACGCGTAAGCGCAGGCTGCGCCAGGCGCTCTCCCCCTCTTGGCCAGCTCTTCA 303  
Qy 241 GCTAGCGAGTTGCTGCGCCCTTCTCGAAGGCGACATATTAACCAAAATGACTTCCCA 300  
Db 304 GCTAGCGAGTTGCTGCGCCCTTCTCGAAGGCGACATATTAACCAAAATGACTTCCCA 363  
Qy 301 GACGCTGACCTCATCGAGGCGCAACTCTGTGGCGGCATGAGATGGGCGGGACATTACC 360  
Db 364 GACGCTGACCTCATCGAGGCGCAACTCTGTGGCGGCATGAGATGGGCGGGACATTACC 423  
Qy 361 CGCGTGAATCAGAGAAACAAGTAGTAATCTCGACTCTTTTGAACCCGCTCCGAGCGGAG 420  
Db 424 CGCGTGAATCAGAGAAACAAGTAGTAATCTCGACTCTTTTGAACCCGCTCCGAGCGGAG 483  
Qy 421 GAGGATGACGCGGAAGTGTCCGCTCCGCGGAGAGATCTCGCGGAATTCAGAAATTTCCCA 480  
Db 484 GAGGATGACGCGGAAGTGTCCGCTCCGCGGAGAGATCTCGCGGAATTCAGAAATTTCCCA 543  
Qy 481 CGAGCGATGCGCGCATGCGCAGCGGATTAACAACCTTCGCTGTGAGATCTTGAAG 540  
Db 544 CGAGCGATGCGCGCATGCGCAGCGGATTAACAACCTTCGCTGTGAGATCTTGAAG 603  
Qy 541 GCGCGGAGATGAGTCTCTCTCATGTGTATCATGGGTGCCACTGCGACCTTCTAAGACCCCT 600  
Db 604 GCGCGGAGATGAGTCTCTCTCATGTGTATCATGGGTGCCACTGCGACCTTCTAAGACCCCT 663  
Qy 601 CCTATACACCTTCAACGAGAAAGAGAGAGAGTGTCTGACAGAAATTCACCGGTCTTCT 660  
Db 664 CCTATACACCTTCAACGAGAAAGAGAGAGAGTGTCTGACAGAAATTCACCGGTCTTCT 723  
Qy 661 GCGCTGCGGAGCTTGCACAAGAGGCTTTGTAGTCTCCGACCGGTGCGCGGTGCAAGC 720  
Db 724 GCGCTGCGGAGCTTGCACAAGAGGCTTTGTAGTCTCCGACCGGTGCGCGGTGCAAGC 783  
Qy 721 GCGACGCGCAACCGCCCTCTCTGACCAATCTCTCCAGAGAGCGGAGACAGATCTGACCT 780  
Db 784 GCGACGCGCAACCGCCCTCTCTGACCAATCTCTCCAGAGAGCGGAGACAGATCTGACCT 843  
Qy 781 GAGTCGTATTCCTCATGCGCCCTCTTGAAGGAGGACCGGAGGACCCCGATCTCAGCGAC 840  
Db 844 GAGTCGTATTCCTCATGCGCCCTCTTGAAGGAGGACCGGAGGACCCCGATCTCAGCGAC 903  
Qy 841 GGGCTCTGTGCTTACCGTGAAGTGAAGAGCGCGGTGAGAGCGTGTCTGCTCATGATGCC 900  
Db 904 GGGCTCTGTGCTTACCGTGAAGTGAAGAGCGCGGTGAGAGCGTGTCTGCTCATGATGCC 963  
Qy 901 TACACATGAGACAGGCGCTCTGATCAGCGCATGCGTGGGAGAAAGCAAGCTGCCATC 960  
Db 964 TACACATGAGACAGGCGCTCTGATCAGCGCATGCGTGGGAGAAAGCAAGCTGCCATC 1023  
Qy 961 AACGCGTTGAGCAACTCTTGTGCGGTGACCAACAATGATGCTAGCGTCAACATCCGCG 1020  
Db 1024 AACGCGTTGAGCAACTCTTGTGCGGTGACCAACAATGATGCTAGCGTCAACATCCGCG 1083  
Qy 1021 AGCGCAAGCGCAAGCGG 1035  
Db 1084 AGCGCAAGCGCAAGCGG 1098

RESULT 7  
A28157  
LOCUS A28157 3750 bp DNA linear PAT 04-AUG-1995  
DEFINITION PT-NANBH mRNA fragment from patent GB2239245.  
ACCESSION A28157  
VERSION A28157.1 GI:1248640  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
other sequences: artificial sequences.  
REFERENCE  
1 (bases 1 to 3750)  
AUTHORS  
TITLE Post-transfusal non-A non-B hepatitis viral polypeptides  
JOURNAL Patent: GB 2239245-A 22 26-JUN-1991;

FEATURES  
source  
Location/Qualifiers  
1..3750  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="config formed by cDNA clones from 3' end"  
<1..>3750  
/codon\_start=1  
/transl\_table=11  
/product="PT-NANBH polypeptide"  
/protein\_id="CA01932.1"  
/db\_xref="GI:1248641"  
/translation="MEGVFTGLTVDAHFLEQTRQAGDNPFLVAYQATVCARAQAP  
PSPDMQMLIRLKEPTLRGAPPLRLGAVNENLTLPILKFLIMACGASALEVVTST  
MVLGVYLAALAAICLTGSAVIVIGRIIISGRPIVLPRELVYQEPDMEBCASHLPY  
IEQGMQLEQFOKOKALGLIOTATQAEAAAVVSKRALETTPAKHMMNITSGIOYL  
AGSLTPENPFIASIMFTTASVTPLTOSTLLNLIIGVYAAQALAPSAASAVYAG  
TAGAVGSIIGKVLVLDLAGYAGVAGALVAFVMEGEMSTEDLVNLPAILSPGA  
LVGVVCAAILLRHVQDEGAVQMMNRLIAPASGNHVSPTHYVPSDAAARVQIIS  
DLITTLQLKRLHWMINEDSTPSCSGMWLDVMDWICVPLADPKTMLSILPLPGVP  
PFGCORGYKGVWBGDMQTTGCGAOLIGHVKXGSMRIYVPTKCSNMHGTPLINAY  
TTGCTPSPAPNYSRALMRVAAREYVTVTRGPHVYVTSMTDVKCCQVPAPEFT  
EVDGRLHRYLPACKPLRBERVTQVGLNQYLVSQLPCEBPVYVITSLTDPHSI  
TAETAKRLIARGSPSLSSASQLSAESSVATYITQNDPFDADLIENLIMRBMG3  
DITRVESENKVIILDSFPLAEEDEREVVPABILRKSKEPPAMPAMARDYNPL  
LESWKAPEYVPPVHGCPLPPTKTPPIPPPRKRTVLTSTVSALAEATKAFGS  
EBSAVDSGTATAPDPSDDGAGSDVSYSMPLRBPDPDPLSDMSWTSVEAG  
EDVCCSMSTVMTGALITPCAAEESKLPINALNSILRHHNMVYATTSRGSQOKVY  
TFPDLQIILDDHYQVLKEMKAKASTYVRKALSYEACLTLPHPAKSFRGCAQDVR  
LSSRAINIRSVMEDELLEDTPTIDITIMANEFVCOPEGRKRPALIVPDLGVR  
VCEKMAIYDVSTLPOAVMSSYGFQYSPGQRFVFLVNAWSKTPMGFAVDTGCPDS  
TVTENDIRVESIYQCCDLAPEAROIRLSLERYIIGPLTNSGOMGQYRCHASGV  
LRTSCGNLTTCYLASAAACRAKIODCTMLVCGGLVVICSAQTOEBAASLRVETEA  
MTRYSAPGDPPEYDELITSSSNVSAHDASGRVYVLTFRDP"

ORIGIN  
Query Match 99.1%; Score 1025.4; DB 6; Length 3750;  
Best Local Similarity 99.4%; Pred. No. 2.2e-226;  
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 ACAGAAATGATGGGGTGGGCTGACAGGTACGCTCGGCGGTGCAAACTCTCTCTACGG 60  
Db 1696 ACAGAAATGATGGGGTGGGCTGACAGGTACGCTCGGCGGTGCAAACTCTCTCTACGG 1755  
Qy 61 GAGGAGTCAATTCAGAGTGGGCTCAACCAATCACTGTGGGTGCGAGTCCCATGC 120  
Db 1756 GAGGAGTCAATTCAGAGTGGGCTCAACCAATCACTGTGGGTGCGAGTCCCATGC 1815  
Qy 121 GAGCCGAACCGGATGTGACAGTGTCACTTCACTGCTACCGACCCCTCCCAATCA 180  
Db 1816 GAGCCGAACCGGATGTGACAGTGTCACTTCACTGCTACCGACCCCTCCCAATCA 1875  
Qy 181 GCAAGACAGGCTAAGCGGAGGCTGGCCAGGGGGTCTCCCCCTCTTGGCCAGCTTTCA 240  
Db 1876 GCAAGACAGGCTAAGCGGAGGCTGGCCAGGGGGTCTCCCCCTCTTGGCCAGCTTTCA 1935  
Qy 241 GCTAGCAGTTGTCTGCGCCCTTCTCGAAGGCGCATATCATTAACCAAAATGACTTCCA 300  
Db 1936 GCTAGCAGTTGTCTGCGCCCTTCTCGAAGGCGCATATCATTAACCAAAATGACTTCCA 1995  
Qy 301 GACGCTGACCTCATCGAGGCGCAACTCTGTGCGGCGCATGAGATGGGCGGGGACATTACC 360  
Db 1996 GACGCTGACCTCATCGAGGCGCAACTCTGTGCGGCGCATGAGATGGGCGGGGACATTACC 2055  
Qy 361 CGCGTGAATCAGAGAAACAAGTAGTAATCTCGACTCTTTTGAACCCGCTCCGAGCGGAG 420  
Db 2056 CGCGTGAATCAGAGAAACAAGTAGTAATCTCGACTCTTTTGAACCCGCTCCGAGCGGAG 2115  
Qy 421 GAGGATGACGCGGAAGTGTCCGCTCCGCGGAGAGATCTCGCGGAATTCAGAAATTTCCCA 480  
Db 2116 GAGGATGACGCGGAAGTGTCCGCTCCGCGGAGAGATCTCGCGGAATTCAGAAATTTCCCA 2175  
Qy 481 CGAGGATGCGCGCATGCGCAGCGGATTAACAACCTTCGCTGTGAGATCTTGAAG 540



ORIGIN										
Query Match	99.4%;	Score 1025.4;	DB 6;	Length 3750;						
Best Local Similarity	99.4%;	Pred. No. 2.2e-226;								
Matches 1029;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;						
QY	1	ACAGAGTGAATGGGGGTGGCGGCTGCACAGGTACGCTCTCGGCGTGCAAACCTCTCTCAACG	60							
DB	1696	ACAGAAAGGTATGGGGGTGGCGGCTGCACAGGTACGCTCTCGGCGTGCAAACCTCTCTCAACG	1755							
QY	61	GAGGAGGTGCACATTCGACAGTTCGGGCTCAACCAATCACTGGTTGGGTGGAGCTCCACATGC	120							
DB	1756	GAGGAGGTGCACATTCGACAGTTCGGGCTCAACCAATCACTGGTTGGGTGGAGCTCCACATGC	1815							
QY	121	GAGCCCGCAACCGAGATGATAGCAGTCTCACTTCCATGCTCACCGACCCCTCCACATACAC	180							
DB	1816	GAGCCCGCAACCGAGATGATGACAGTCTCACTTCCATGCTCACCGACCCCTCCACATACAC	1875							
QY	181	GCAGAGACGGGTAAAGCCGAGGCTGGCCAGGGGGTCTCTCCCTCTCTTGGCCAGCTCTTCA	240							
DB	1876	GCAGAGACGGGTAAAGCCGAGGCTGGCCAGGGGGTCTCTCCCTCTCTTGGCCAGCTCTTCA	1935							
QY	241	GCTAGCCCAAGTTGTCTGGCCCTTCTCTCGAAGGCCACATACATTAACCAAAATGACTTCCCA	300							
DB	1936	GCTAGCCCAAGTTGTCTGGCCCTTCTCTCGAAGGCCACATACATTAACCAAAATGACTTCCCA	1995							
QY	301	GACGCTGACCTCATATGAGGCGCAACCTCCTGGGCGGATGAGATGGGCGGGGACATTAC	360							
DB	1996	GACGCTGACCTCATATGAGGCGCAACCTCCTGGGCGGATGAGATGGGCGGGGACATTAC	2055							
QY	361	CGCGTGAAGTCAAGAAACAAGTAGTAATCTCTGACCTTTTCGACCCGCTCCAGCGGAG	420							
DB	2056	CGCGTGAAGTCAAGAAACAAGTAGTAATCTCTGACCTTTTCGACCCGCTCCAGCGGAG	2115							
QY	421	GAGGATGAGCGGGAGAGTGTCCGTCCGCGCGGAGATCTCTGGGAGAAATCCAGAAATTTCCCA	480							
DB	2116	GAGGATGAGCGGGAGAGTGTCCGTCCGCGCGGAGATCTCTGGGAGAAATCCAGAAATTTCCCA	2175							
QY	481	CCAGCGAATGCCCGATGGGGACCGCGCGGAGTTACCAACCTCCGCTCGTAGAGTCTCTGAG	540							
DB	2176	CCAGCGAATGCCCGATGGGGACCGCGCGGAGTTACCAACCTCCGCTCGTAGAGTCTCTGAG	2235							
QY	541	GCCTCCGACCTACGTCCTCTCCAGTGGTACATGGGTGCCCATGTCACCTACTAAGACCCCT	600							
DB	2236	GCCTCCGACCTACGTCCTCTCCAGTGGTACATGGGTGCCCATGTCACCTACTAAGACCCCT	2295							
QY	601	CTTATACCACTCCACGAGAGAAAGAGACAGTTGTTCTGACAGATTCACCGTGTCTTCT	660							
DB	2296	CTTATACCACTCCACGAGAGAAAGAGACAGTTGTTCTGACAGATTCACCGTGTCTTCT	2355							
QY	661	GGCCTGGCGGAGCTTTGGCCAAAGGCTTTTGGTAGTCTCGGACCGTGGGCGGTGAGACGC	720							
DB	2356	GGCCTGGCGGAGCTTTGGCCAAAGGCTTTGAGTAGTCTCCGAACCTGTGGCGGTGAGACGC	2415							
QY	721	GGCAGCGCAACCGCCCTCTCTGACCAATCTCTCCGACGACGAGCGAGACAGATCTGACGTT	780							
DB	2416	GGCAGCGCAACCGCCCTCTCTGACCAACCTCTCCGACGACGAGCGAGATCTGACGTT	2475							
QY	781	GAGTCGTATTTCTTCATGCCCCCTTTGAGGGGGAGGCGGGGAGCCCCGATTTCAAGCGAC	840							
DB	2476	GAGTCGTATTTCTTCATGCCCCCTTTGAGGGGGAGGCGGGGAGCCCCGATTTCAAGCGAC	2535							

QY 841 GGGCTTGGTCTACCGTGAAGAGAGCGCGGTGAAGACGTCTGTCTGATGTC 900  
|  
|  
|  
Db 2536 GGGCTTGGTCTACCGTGAAGAGAGCGCGGTGAAGACGTCTGTCTGATGTC 2595  
|  
|  
|  
QY 901 TACACATGACAGCGCGCTCTGATCACGCCATGCGTCCGAGAAAGAAAGCTGCCATC 960  
|  
|  
|  
|  
Db 2596 TACACATGACAGCGCGCTCTGATCACGCCATGCGTCCGAGAAAGAAAGCTGCCATC 2655  
|  
|  
|  
QY 961 AACGGGTGAGCAACTTGTGCGGTGACCAACAATGATGCTACGCTACCAATCCGCG 1020  
|  
|  
|  
|  
Db 2656 AACGGGTGAGCAACTTGTGCGGTGACCAACAATGATGCTACGCTACCAATCCGCG 2715  
|  
|  
|  
QY 1021 AACGCAAGCCAGCGG 1035  
|  
|  
|  
Db 2716 AGCGCAAGCCAGCGG 2730  
|  
|  
|  
RESULT 9  
ARI44051 3750 bp DNA linear PAT 08-AUG-2001  
LOCUS Sequence 22 from patent US 6210675.  
ACCESSION ARI44051  
VERSION ARI44051.1 GI:15105918  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 3750)  
AUTHORS Highfield,P.Edmund., Rodgers,B.Colin., Tedder,R.Seton. and  
Barbara,J.Anthony.James.  
TITLE PT-NANB hepatitis polyproteinides  
JOURNAL Patent: US 6210675-A 22 03-APR-2001;  
FEATURES Location/Qualifiers  
source 1..3750  
/organism="unknown"  
/mol\_type="unasigned DNA"

Query Match 99.1%; Score 1025.4; DB 6; Length 3750;  
Best Local Similarity 99.4%; Pred. No. 2.2e-226;  
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ACAGAGTGTGATGGGTGCGGCTGACAGGTACGCTCCGCGGTGCAAACTCTCTACGG 60  
|  
|  
|  
Db 1696 ACAGAGTGTGATGGGTGCGGCTGACAGGTACGCTCCGCGGTGCAAACTCTCTACGG 1755  
|  
|  
|  
QY 61 GAGGAGGTCACTTCAGGTCCGGCTCAACCAATACCTGTTGGTGGAGCTCCCATC 120  
|  
|  
|  
Db 1756 GAGGAGGTCACTTCAGGTCCGGCTCAACCAATACCTGTTGGTGGAGCTCCCATC 1815  
|  
|  
|  
QY 121 GAGCCCGAAGCCGATGTAGCAAGTCACTTCATGCTCAACGACCCCTCCCAATCA 180  
|  
|  
|  
Db 1816 GAGCCCGAAGCCGATGTAGCAAGTCACTTCATGCTCAACGACCCCTCCCAATCA 1875  
|  
|  
|  
QY 181 GAGAGACCGCTTAAGCGAGCTGCGCAAGGAGTCTCCCTCTCTTGGCCAGCTCTTA 240  
|  
|  
|  
Db 1876 GAGAGACCGCTTAAGCGAGCTGCGCAAGGAGTCTCCCTCTCTTGGCCAGCTCTTA 1935  
|  
|  
|  
QY 241 GCTAGCCAGTTGTCTGCGCTTCTCTGAAAGCGACATACATTAACCAAAATGACTTCCA 300  
|  
|  
|  
Db 1936 GCTAGCCAGTTGTCTGCGCTTCTCTGAAAGCGACATACATTAACCAAAATGACTTCCA 1995  
|  
|  
|  
QY 301 GAGCGTCACTCATGAGGCGCAACCTCCGTGGCGGCGAGATAGGCGGGGACATTAAC 360  
|  
|  
|  
Db 1996 GAGCGTCACTCATGAGGCGCAACCTCCGTGGCGGCGAGATAGGCGGGGACATTAAC 2055  
|  
|  
|  
QY 361 CGCGTGAATCAGAGAAACAAGTAGTAATCTGAACTCTTGAACCCGCTCCGAGGAG 420  
|  
|  
|  
Db 2056 CGCGTGAATCAGAGAAACAAGTAGTAATCTGAACTCTTGAACCCGCTCCGAGGAG 2115  
|  
|  
|  
QY 421 GAGGATGACCGGAAAGTGTCCGTCCCGGCGAGATCTCCGAAATCCAAGAATTTCCA 480  
|  
|  
|

Db 2116 GAGGATGACCGGAAAGTGTCCGTCCCGCGGAGATCTTCGAGAAATCCAAAAATTTCCA 2175  
|  
|  
|  
QY 481 CGAGGATGACCGGAAAGTGTCCGTCCCGCGGAGATCTTCGAGAAATCCAAAAATTTCCA 540  
|  
|  
|  
Db 2176 CGAGGATGACCGGAAAGTGTCCGTCCCGCGGAGATCTTCGAGAAATCCAAAAATTTCCA 2235  
|  
|  
|  
QY 541 GCCCGGACCTACGTCCTCCAGTGTATCATGGGTGCCACTGACCACTTAAGAACCCT 600  
|  
|  
|  
Db 2236 GCCCGGACCTACGTCCTCCAGTGTATCATGGGTGCCACTGACCACTTAAGAACCCT 2295  
|  
|  
|  
QY 601 CCTATACCACTTCCAGGAGAAAGAGACAGTTGTTCTGAACAATCCACCTGCTTCT 660  
|  
|  
|  
Db 2296 CCTATACCACTTCCAGGAGAAAGAGACAGTTGTTCTGAACAATCCACCTGCTTCT 2355  
|  
|  
|  
QY 661 GCCCTGGGAGCTTGCACAAAGCTTTGTGATGCTCCGAGCCGTGGCCGTGAGACG 720  
|  
|  
|  
Db 2356 GCCCTGGGAGCTTGCACAAAGCTTTGTGATGCTCCGAGCCGTGGCCGTGAGACG 2415  
|  
|  
|  
QY 721 GGCACGGCAACCGCCCTCTGACCAATCTCCGACGACGCGGAGACAGATCTGACGTT 780  
|  
|  
|  
Db 2416 GGCACGGCAACCGCCCTCTGACCAATCTCCGACGACGCGGAGACAGATCTGACGTT 2475  
|  
|  
|  
QY 781 GAGTCTATTCTCTCATGTCCTCCCTTTGAGGGGAGACCTGGGAGACCCGATCTGACGAC 840  
|  
|  
|  
Db 2476 GAGTCTATTCTCTCATGTCCTCCCTTTGAGGGGAGACCTGGGAGACCCGATCTGACGAC 2535  
|  
|  
|  
QY 841 GGGCTTGGTCTACCGTGAAGAGAGCGGTGACAGCTGCTGCTGATGTC 900  
|  
|  
|  
Db 2536 GGGCTTGGTCTACCGTGAAGAGAGCGGTGACAGCTGCTGCTGATGTC 2595  
|  
|  
|  
QY 901 TACACATGACAGCGCGCTCTGATCACGCCATGCGTCCGAGAAAGAAAGCTGCCATC 960  
|  
|  
|  
Db 2596 TACACATGACAGCGCGCTCTGATCACGCCATGCGTCCGAGAAAGAAAGCTGCCATC 2655  
|  
|  
|  
QY 961 AACGGGTGAGCAACTTGTGCGGTGACCAACAATGATGCTACGCTACCAATCCGCG 1020  
|  
|  
|  
Db 2656 AACGGGTGAGCAACTTGTGCGGTGACCAACAATGATGCTACGCTACCAATCCGCG 2715  
|  
|  
|  
QY 1021 AGCGCAAGCCAGCGG 1035  
|  
|  
|  
Db 2716 AGCGCAAGCCAGCGG 2730  
|  
|  
|  
RESULT 10  
A76577 7065 bp DNA linear PAT 19-OCT-1999  
LOCUS Sequence 9 from Patent W09317110.  
ACCESSION A76577  
VERSION A76577.1 GI:6088488  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 7065)  
AUTHORS Rodgers,B.C. and Parker,D.  
TITLE A RECOMBINANT HEPATITIS C VIRUS POLYPEPTIDE  
JOURNAL Patent: WO 9317110-A 9 02-SEP-1993;  
FEATURES Location/Qualifiers  
source 1..7065  
/organism="unidentified"  
/mol\_type="unasigned DNA"  
/db\_xref="taxon:32644"  
<1..7065  
/note="Bncodes putative PT-NANBH non-structural proteins"  
/codon\_start=1  
/product="PT-NANBH coding sequence"  
/protein\_id="CAB58625.1"  
/db\_xref="GI:6088489"  
/translation="SCGAVPFGIMPLTSLPYRVPLAKLIMWQYPTREAHLOW  
VPLNVRGRDAIITLACVHBEIFIDISKLLILGLMLLQIGITRVYFVBAQGL  
IRACMLVRKVGHHVQMALMLKLAALIGTYVYDHLTFLRDAHAGLDAVAEPPVE  
SDMETKVIWGTDTAACGDITLIGLPVBARBDEIFLDPADSLBGOGRLLAPITAYSO

OTRGLGCIITSLTGRDKNOVEGEVVOVNSTATOSPLATCVNGCMTVYHGAGSKTLG  
PGRPVOTMYNDODLVGWPAPRGASLITPCGSSDILYTHMHAVIIPRRGDSRG  
SILSPRIYIUNGSSGGPILCPBGHVGIFRAVCRGVAKANDPIPVSEMETTVASR  
VFDNSPPRAVPOSFVVAHDAHPTGSGKSTRVPAAYAAQCTKVLNVPSPVMA/LGGA  
YMSKAGVDPNIRSGVRIITVGAPIITYSYGKFLADGSGGAVDIIIDCECHSTST  
SILIGCTVDOAETAGARLVLATATPPGVSPVPHNIEGALSTNIEJLPEFKAPI  
ERTKGRHLIFCHSKKDEBLAKLVGLGNVAVAYATVPSAGDVVAVATLAL  
MGFTGDPDSVIDCMTCTVOTVDPISLDPITTEITTVPODVAVSROBRGTGGRGI  
YRPAVGERPSGMPDSSVCECTDAGCANTELTPAETSVRIATYIANTPGLPVODHLE  
FNEBFTGLIHDVDAFLSOTKQNDNPIVATQATCAQAPSPSMDM/CLLRL  
KPTLRGPTLVRIGAVONEVTLTHPIKFKVIMACMADIEVNTVSTVLVGVVLAAL  
YCLTTSVVIYGR11ILSGRPAIVPDRLEVIOEFDEMEBEGSHLYLGEOLAEQKO  
KALGLTOTAKOEAAPVYESKRALETEFMAKMMNFISIGIYLAGSLTPENPPIA  
SILMAFTASVTSPLTOSTLILNIGGMVAOLAIPSAASFGAGVAGVAGVIGIGK  
VLVDILAGVAGVAGVAGVAFKMSGEMPTEDLVNLLPALSGALVGVCAITIRK  
HVGEGVAVQMMNRLIAFASRGNHVSPTHVPESDAARTQILSDTLTQILKRLHO  
WINECSTPCSGSLRDVMDICTVLADPFTWLOSILPLPVPFSPSCORGKXKAY  
GDGIMQTCSCGKQITGVHNGSMRIVGPTCSNMHMGEPPIINAYTTGCTPSPANV  
SRLAMVAAEEYEVTRVGDPHYVTSMITDNVCKPOVPAPEFTVEDVGRHLRVAPA  
CKPLAEEVTFVQGLNOYLVGSQLPCEPEPDVAIVLSMLTPBHTATRAKRLAAGS  
PSBLASSASQLSAPSSKATYITONDPPBDLIEANULRHEMGGITTRYESNKYIT  
LOSFPDLRAEBDEREVSPAEILRKSKKEPPAPAMARPDVNPPLLESWAPDYVAP  
VHGCPLEPTKTPPIPPRRKRTVLTSTVSSALAEIATAFSGSDVAVSCSWSYMT  
POPSBDGAGSDVESYSSMPLEGPDPDLSDGSMSTVSEAGSDVVCSSWSYMT  
GALITPCABESKLPINALSNSILRHHNVYATTSRSASOROKKTPEDRLQILDHXY  
DULKEMKAKASTYKALISVEACIKLTPHSASKRGYKAKYRNLSKXINIRBYM  
EDLEBETRIPIITTTAKNEVFCVOPBERGKRPALIVPDLVRCERKALYDVST  
LPOAVNGSSYGFQYSPQARVFLVNAWKSKTPMGCPAYDRCSVTENDRIVEEVK  
IKALISVEACIKLTPHSASKRGYKADVNLSKALINHRIMEDLLEBETRIPIIT  
TAKNEVFCVOPBERGKRPALIVPDLGVEKSKALIDVSTLPOAVNGSSYGOY  
SPOGRVEFLVNAWKSKTKPMGFAYDRCPSTVENDIRSEIYOCCEADPARARAI  
RSLTERLYIGPLTNSKGNOCYRCRCAAGVLTSCGNTLTCLYKASACRAKALODC  
THIVCDDLVIVICESAGTQDASLAVTFBAMTRYSAFPEDPPOPEYDELTJCSGN  
VSVADHDSGRVYLTTRDPTTLARAMEARTPVMNSGIMIAPIWALMAIIMLT  
HEFESILAOBOLKALDCOYGAYSIEPDLPOIIORLHGSLASISHSPEIRNV  
ASCLRIQVPLRVRHRAVARAKLISOGGRAICRGYIFMAVATKLTPIIPAS  
RLDLSGMFVAGYSGGDIYHSVBARPRMFWMLLSLVGVIYLLPNR"  
7063 7065  
/function="Stop codon"

## ORIGIN

Query Match 99.1%; Score 1025.4; DB 6; Length 7065;  
Best Local Similarity 99.4%; Pred. No. 2.1e-226;  
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ACAGAGTGGATGGGCTGCGCTGCAAGGATGCTTCGCGGCTGCAAACTCTCTTAACG 60  
DB 3919 ACAGAGTGGATGGGCTGCGCTGCAAGGATGCTTCGCGGCTGCAAACTCTCTTAACG 3978  
QY 61 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATTAAGTGGGCTGCGAGTCCCATGCG 120  
DB 3979 GAGAGAGTCAATTCAGAGTGGGCTCAACCAATTAAGTGGGCTGCGAGTCCCATGCG 4038  
QY 121 GAGCCCGAACCAGATGAGAGTGCATCTTCATGCTCCAGACCCCTCCCATCATACA 180  
DB 4039 GAGCCCGAACCAGATGAGAGTGCATCTTCATGCTCCAGACCCCTCCCATCATACA 4098  
QY 181 GCGAGAGCGGCTAAGCGCAAGCTGCGCAAGGAGGTCTCCCTCTCTTGGCCAGCTCTTCA 240  
DB 4099 GCGAGAGCGGCTAAGCGCAAGCTGCGCAAGGAGGTCTCCCTCTCTTGGCCAGCTCTTCA 4158  
QY 241 GCTAGCGAGTGTCTGCGCCCTTCTCGAAGGCGACATACATTAACCAAAATGACTTCCCA 300  
DB 4159 GCTAGCGAGTGTCTGCGCCCTTCTCGAAGGCGACATACATTAACCAAAATGACTTCCCA 4218  
QY 301 GACGCTGACCTCATCGAGGCGCAACTCTTGGCGGCGATAGATGGCGGGGACATTACG 360  
DB 4219 GACGCTGACCTCATCGAGGCGCAACTCTTGGCGGCGATAGATGGCGGGGACATTACG 4278  
QY 361 CGGCTGAGAGTCAAGAAACAAGTAGTAATCTTGAAGCTCTTTCGACCGGCTCGAGCGGAG 420  
DB 4279 CGGCTGAGAGTCAAGAAACAAGTAGTAATCTTGAAGCTCTTTCGACCGGCTCGAGCGGAG 4338

QY 421 GAGATAGACGGGAAAGTGTCCGTCGGCGGAGATCCTCGGAAATCCAAGAAATTCCTCA 480  
DB 4339 GAGATAGACGGGAAAGTGTCCGTCGGCGGAGATCCTCGGAAATCCAAGAAATTCCTCA 4398  
QY 481 CCAGCGATGCCCGATAGGCGACGCCCGGATTAACAACCTTCGCTGAGATCTTGAG 540  
DB 4399 CCAGCGATGCCCGATAGGCGACGCCCGGATTAACAACCTTCGCTGAGATCTTGAG 4458  
QY 541 GCCCGGATTAAGTCCCTCCAGTGTGATCACTGGGTGCCCATCTGCACCTTAAGACCCCT 600  
DB 4459 GCCCGGATTAAGTCCCTCCAGTGTGATCACTGGGTGCCCATCTGCACCTTAAGACCCCT 4518  
QY 601 CCTATACACCTTCCACGAGAAAGAGACAGTGTCTGACAGAAATCCACCGGTCTTCT 660  
DB 4519 CCTATACACCTTCCACGAGAAAGAGACAGTGTCTGACAGAAATCCACCGGTCTTCT 4578  
QY 661 GCCCTGCGGAGCTTGCACAAAGGCTTTGTAGTCTCGGACCGTCCGCTGACAG 720  
DB 4579 GCCCTGCGGAGCTTGCACAAAGGCTTTGTAGTCTCGGACCGTCCGCTGACAG 4638  
QY 721 GCGACGCGCAACCGCCCTCTGACCAATCTCTCGAGACGCGGAGACAGATCTGACGTT 780  
DB 4639 GCGACGCGCAACCGCCCTCTGACCAATCTCTCGAGACGCGGAGACAGATCTGACGTT 4698  
QY 781 GAGTGTATTCTTCATGCCCCCTTGAAGGGGAGCCGGGGACCCCGATCTCAGCGAC 840  
DB 4699 GAGTGTATTCTTCATGCCCCCTTGAAGGGGAGCCGGGGACCCCGATCTCAGCGAC 4758  
QY 841 GGGTCTTGTGCTTACCGTGAAGTGAAGAGCGCGGTGAGACGTCGTCTGCTGATGTCC 900  
DB 4759 GGGTCTTGTGCTTACCGTGAAGTGAAGAGCGCGGTGAGACGTCGTCTGCTGATGTCC 4818  
QY 901 TACACATGACACAGCGCTCTGATCAGCCATGCGCTGCGGAGAAACAGAGTCCCATC 960  
DB 4819 TACACATGACACAGCGCTCTGATCAGCCATGCGCTGCGGAGAAACAGAGTCCCATC 4878  
QY 961 AACGCGTTAGCAACTCTTGTGCGGTGACCAACCAATGATGATCTGACATCCCGG 1020  
DB 4879 AACGCGTTAGCAACTCTTGTGCGGTGACCAACCAATGATGATCTGACATCCCGG 4938  
QY 1021 AGCGCAAGCCAGCGG 1035  
DB 4939 AGCGCAAGCCAGCGG 4953

RESULT 11  
AF313916  
LOCUS AF313916 9359 bp mRNA linear VRL 01-JAN-2002  
DEFINITION Hepatitis C virus polyprotein mRNA, partial cds.  
ACCESSION AF313916  
VERSION AF313916.1 GI:18027684  
KEYWORDS  
SOURCE Hepatitis C virus  
ORGANISM Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.  
1 (bases 1 to 9359)  
Fanning,J., Itakura,J., Nagayama,K. and Enomoto,N.  
Characteristics of Hepatitis C viral genome associated with disease  
progression in a homogeneous patient population  
Unpublished  
JOURNAL  
2 (bases 1 to 9359)  
Fanning,J., Itakura,J., Nagayama,K. and Enomoto,N.  
Direct Submision  
Submitted (17-OCT-2000) Medicine, National University of Ireland,  
Cork, Hepatitis C Unit, Clinical Sciences Building, Cork University  
Hospital, Cork, Ireland  
FEATURES  
source  
1..9359  
/organism="Hepatitis C virus"  
/mol\_type="mRNA"  
/db\_xref="taxon:11103"  
/note="Isolated from source plasma identified as the

single source of the iatrogenic infection of Rheus negative women in Ireland from May 1977-November 1978; exposure was through HCV 1b contaminated anti-D immunoglobulin"

5'UTR  
CDS

330...>9359  
/codon\_start=1  
/product="polypotein"  
/protein\_id="AAU55821.1"  
/db\_xref="GI:16027685"

/translation="MSTNPKPQRTKNTNRPRDVKPRGGQIVGVYLLPRRPRRL  
GRATRKTSERQPRGRQPIPKARBERTAQPGYEMPLVGNAGAGMGLVLRG  
SRSPGPTDPRRRSRNLKVIDTLTCGGFADMGYIPLVAGARMLAHGVRLD  
GNVATGNLPGCSFSIPLIALISCLTIPASIAEVNRSQVYTNDSIAVYAAD  
MIMHPGCVPCVRENNSRCVVALPPLAARNSIPLTIRHVTLVLDGASAFYAG  
VVDLGSVFLVQLFTFSPRHEWQDNCSLYPGHVSQHRMAMMMNMTTLVY  
SGLATIPQAVNDMAHGVYAGLAIYISMAQNAKVLVLLPFGVDGETTGAQ  
AHTRGATLTFSPSQNTQLINTGSMHINTALNCDSLNTGFLALFVAHRNAS  
GCPERMASCRPIDKPAQCGPIAYAKPLSLDKPCMHYAPQCIIVASQVCGVYC  
FTSPVVVGTDRFGVPTYSNGENEDVLLNTRPBRNMGCTMANGTPTKCG  
PECDIGVGNNDLTCPTDFRKHPEATYKCGSPMLPRCWADVPLRMHPCVNF  
TIFKRVYVGVGVEHLNMAQWTRGBRCDLDRDRSELPLISTEMOVLPSTTLL  
PALSTGLIHLNRNIDVQYLYIGISAVSYAIKREYVLLFTPLADARYCCLMMLL  
IVQAEALENLVNNAASVAGTHGISLPLFPCAAWYIKGRVLPQAAAYAFGVWDL  
LLALPBRVYAMDREMAASCGAVFIPLALISLPHYKVLARLIMVLYPTREAH  
LQVWIPPLVNRGDAIILMCAIHPELIPTITLILMIGPLVLOAGITVPEVR  
AHGIRACVLRVAGHYOMALMRALGTYYVDLITPLDHAHSELRDLAVAE  
PVFSDMETKLIITWGNADTAACGDIILGIPVSRGRREILLAGNDSLEGGMLA  
PITAYOORGLGCTIYSLTRGRDNQVYGEVYSTAOSTLACVNGCTYTHAGTK  
TLAQRGPTIQTNTVDDVGLMLAPGKASLPTCCGSDILVTRHADVLVARRG  
DSRGLSLSPVSYLKSGSGPLCPGSAVGIIRAAQCTRVAAAVFVPEVSMET  
NMSPVFTDNSSPPAVPOTFQVNAHAPTSQSGSTVPAAYAGQKAVLVNPSVATL  
GFGAMSKAHGVDPNIRNGARTITGAPITYSTYKGLADGSGAAVDIICDEKS  
TDSITLIGTGLDQAEATGARLVVLAATPGSVTPHPNTEVALSNIGIIPCHG  
AIPITIGSRHLIRCHSRKCDLAAKLSGLANAVAYRDLDSVPTSGDVVVA  
TDALMTGTDFDSIDCVTCQVYVSDPTFTLETIVQDAVYRDLDSVPTSGDVVVA  
RTGIYRPTVPRPSGMDSVLCEYDAGCAMYELTPAETSVLRALANTPGLPVCQ  
DHLFEWEGFTGLTHIDAFLSQTKQAGNPNYLAQVATYCARAQAPSPMDQMK  
LIRLKPETHGFTPLIRLGAQVQNEVILHPIKPIFMAQASALEVATSTWVVGVA  
ALAAVCLTGSVIVIRIILSGKPAIVPRREVLYOOPMEMECALHLYIRGOMLAE  
OPKQKAGLLOTRTAQABAAVYVSKRALESFAKMMNITSGIYLAISTIPGN  
PAIASLMAFTASITPLTQHTLNFNIGLGVAAQLAPPASAVVAGIAGAAV  
IGKXVLVILAGYAGVAGALVAFKVMGEMSTEDLVNLLPAIISPALVGVVCA  
ILRHVYGBEGAVQMMNRLIAPASRGNYSPHYVPESDAARVYIISLITLTLK  
RLHWINEDCTSPGSMRLQVDMWICVLTDFPTKMOSKILPRLPGVPSGCKG  
GVMRGDMQCTPCGAOITGHVKNKSRIVGRTCSMTGHEPTLINTYTPGCSPP  
APNYSBALMRVTAEEYLYEYQYGDPHYVGTMTDVKPCQVPAEPTFEBDVGLR  
YAPCKPLAREVITFQVLANQYLVESQLECEPEVAVLTSMLTDPHSITAEAKRL  
AKGSPSLASSASQLSAPSLKATCTTRHSDPADLIEANILMRQEMGNITRVSEN  
RVVILDSFEPPLAEEDEREVSYPALILRSRKFPRAMPIMADPNPPLVESKMDPY  
VPPVHGCPLPTKAPLIPPRKRIVLTSTESVSALEAKTKFSGSESAWQGT  
ATAPDPOSDDDAGSDVESYAMPLEGBEPDPLISGDSYVEEASBPVCCSMS  
YTMGALITPCSAEESKLPINALSNLSILRHNMVYATTSRSASQKQVCTPRLOYLD  
DHYRVLKMKKASTVAKLLPVEACKLTPPHSAKSLFEGAKDQVAKVCTPRLOYLD  
LSVMKDLLEDTEPIDTITIMANEVFCVQPEKGRKPARLVFPLGAVRSEKALVY  
VSTLPQAVMGSSYGFQYSPQORVEFLVAMKCKKPNFADTFCSTVYESIRY  
RESIYOCCLAPBARQAIRSLTERLYIGGPLTNSKQMGXSVRCASTGLTSCNLT  
TCYLKASACRAKLODCTMLVCGDVLVICSAGTIDBAISIRVETAMRYSPG  
DPRREYDLLETSCSSNVSVNADASGRVYVLTDPPTTTLARAMWERTARHPVNSL  
GNIMYAPETLARMLHTHPSITLAQOBLKALDCOTYGACTSITLDPITDRLH  
GLSATSLSYSPGEINRVAACTRLKGLVPLRWRHRASVRKXKLSQGRALANCKLY  
FNMAVRTKXKLPIDPAASQLDLSWFWAGYSGSDIYHLSIRBRPRWFMLCILLISGV  
GIYLLPNR"

ORIGIN

Query Match 87.0%; Score 900.6; DB 14; Length 9359;  
Best Local Similarity 91.9%; Pred. No. 1.4e-197;  
Matches 951; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 1 ACAGAGTGTGATGGGTGCGCTGCAAGGTACGCTCGCGGCTGCAAACTCTCTACGG 60  
DB 6696 ACGAGGTGATGGGTGCGCTGCAAGGTACGCTCGCGGCTGCAAACTCTCTACGG 6755

QY 61 GAGAGGTCAATTCAGAGTCGAGCTCAACCAATACCTGTTGGGTGCGAGCTCCATGC 120  
DB 6756 GAGAGGTCAATTCAGAGTCGAGCTCAACCAATACCTGTTGGGTGCGAGCTCCATGT 6815  
QY 121 GAGCCGGAACCGGATGTAGCAGTGTCTCATCTTCATGCTCAACCGACCTCCATCA 180  
DB 6816 GAGCCGGAACCGGATGTAGCAGTGTCTCATCTTCATGCTCAACCGACCTCCATYCA 6875  
QY 181 GAGAGGAGGCGTAAAGCGGAGCTGGCGAGGGGGTCTCCCGCTCTTGGGAGCTTCA 240  
DB 6876 GAGAGGAGGCGTAAAGCGGAGCTGGCGAGGGGGTCTCCCGCTCTTGGGAGCTTCA 6935  
QY 241 GCTAGCCAGTGTCTGAGCTCTTCTCGAAGCGCATATCATTAACCAATGACTTCCA 300  
DB 6936 GCTAGCCAGTGTCTGAGCTCTTCTCGAAGCGCATATCATTAACCGATGATGCCCA 6995  
QY 301 GACGCTGACTCATGAGGCGCACTCTGTGGCGGCGATGAGATGGCGGGGACATTAAC 360  
DB 6996 GATGCTGACTCATGAGGCGCACTCTGTGGCGGCGGATGAGATGGCGGGGACATCAAC 7055  
QY 361 GCGGTGAGTCAAGAAAGAAAGTGTATCTCGGACTCTTCAACCGCTCGAGCGGAG 420  
DB 7056 GCGGTGAGTCAAGAAAGAAAGTGTATCTCGGACTCTTCAACCGCTCGAGCGGAG 7115  
QY 421 GAGATGAGCGGGAAGTGTCCGTCCGCGCGAGATCTCCGGAATCCAGAAATTTCCA 480  
DB 7116 GAGATGAGCGGGAAGTGTCCGTCCGCGCGAGATCTCCGGAATCCAGAAATTTCC 7175  
QY 481 CAGGATGATCCCGGATGGGCGACGCGCGGATTAACAACCTCGCTGTGAGATCTTGAA 540  
DB 7176 CAGGATGATCCCGGATGGGCGACGCGCGGATTAACAACCTCGCTGTGAGATCTTGAA 7235  
QY 541 GCCCGGACTGAGTCCCTCCAGTGGTACATGAGGTGCCCACTCACTTAAGACCCCT 600  
DB 7236 GACCCGACTGAGTCCCTCCAGTGGTACATGAGGTGCCCACTCACTTAAGACCCCT 7295  
QY 601 CCTATACCACTCAACGAGAAAGAGCAAGTTGTTCAGACAGATCCACCGTCTTCT 660  
DB 7296 CCAATACCACTCAACGAGAAAGAGCAAGTTGTTCAGACAGATCCACCGTCTTCT 7355  
QY 661 GCCCTGGCGAGCTTGCACAAGAGCTTTGTGCTCGGACCGTCCGCGTGCACGC 720  
DB 7356 GCTTGGCTGAGCTCGCACAAAGACCTTCGAGCTCCGAGTGTGTGCGCGACGC 7415  
QY 721 GGCACGGCAACCGCCCTCCGACCAATCTCCGACGAGCGGAGGAGATTCGAGCTT 780  
DB 7416 GGCACGGCAACCGCCCTCCGACCAATCTCCGACGAGCGGAGGAGATTCGAGCTT 7475  
QY 781 GAGTCGATTCCTCATGCCCCCTTGAAGGGAGGAGCCGAGGAGACCCGATCTGACGAC 840  
DB 7476 GAGTCGATTCCTCATGCCCCCTTGAAGGGAGGAGCCGAGGAGATTCGAGCTT 7535  
QY 841 GGGTCTTGTCTACCGTGAAGTGAAGAGCGGTGAGGACGTCGTCTGCTCGATGTCC 900  
DB 7536 GGGTCTTGTCTACCGTGAAGTGAAGAGCGGTGAGGACGTCGTCTGCTCGATGTCC 7595  
QY 901 TACCATGAGCAGGCGCTGATCAACGCAATGCGGTGAGGAGAAAGCAAGCTGCCATC 960  
DB 7596 TACCATGAGCAGGCGCTGATCAACGCAATGCGGTGAGGAGAAAGCAAGCTGCCATC 7655  
QY 961 AAGCGTTGAGCAACTTTTCTGCTGCTACCAACAATGATCTACGCTACCAATCCGC 1020  
DB 7656 AAGCGTTGAGCAACTTTTCTGCTGCTACCAACAATGATCTACGCTACCAATCCGC 7715  
QY 1021 AGCGCAAGCCAGCGG 1035  
DB 7716 AGCGCAAGCCAGCGG 7730

RESULT 12  
AF208024 AF208024 9373 bp RNA linear VRL 30-MAR-2000  
DEFINITION Hepatitis C virus strain MD34, complete genome.



Qy 781 GAGTCGATTCCTCCATGCCCCCTTGAAGGGGAGCCGGGAGACCCCGATCTCAGCGAC 840  
Db 7470 GAGTCGATTCCTCCATGCCCCCTTGAAGGGGAGCCGGGAGATCCGATCTCAGCGAC 7529  
Qy 841 GGGTCTTGCTTACCGTGAAGTGAAGAGCGCGGTGAAGAACTGCTGCTGCTCGATGTC 900  
Db 7530 GGGTCTTGCTTACCGTGAAGTGAAGAGCGCGGTGAAGAACTGCTGCTGCTCGATGTC 7589  
Qy 901 TACACATGAGACAGCGCTCTGATCAGCCCATGCGCTGCGGAGAGAAAGAGTGCCTATC 960  
Db 7590 TACACATGAGACAGCGCGCTCTGATCAGCCCATGCGCTGCGGAGAGAAAGAGTGCCTATC 7649  
Qy 961 AACGCGTTGAGCAACTTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Db 7650 AATGCGTTGAGCAACTTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7709  
Qy 1021 AACGCGAGCGACCGG 1035  
Db 7710 AACGCGAGCGACCGG 7724

RESULT 13  
AP007753  
LOCUS AP007753 9374 bp RNA linear VRL 27-APR-2000  
DEFINITION Hepatitis C virus strain MD12 complete genome.  
ACCESSION AF207753  
VERSION AF207753.1 GI:7650223  
KEYWORDS  
SOURCE  
ORGANISM Hepatitis C virus  
Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.  
1 (bases 1 to 9374)  
Nagayama, K., Kurotaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and  
Sato, C.  
Characteristics of hepatitis C viral genome associated with disease  
progression  
Unpublished  
2 (bases 1 to 9374)  
Nagayama, K., Kurotaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and  
Sato, C.  
Direct Submission  
Submitted (23-NOV-1999) Second Department of Internal Medicine,  
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,  
Tokyo 113-8519, Japan  
LOCATION/Qualifiers  
1. 9374  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/strain="MD12"  
/db\_xref="taxon:11103"  
/note="Isolated from a patient with persistently-normal  
ALT values (ASC-2)"  
330..9362  
/product="polyprotein"  
/protein\_id="AAF65943.1"  
/db\_xref="GI:7650224"  
/translation="MSTNPKPQKTKKTNRRPQVYKPPGGQIVGVYLLPRRPRLL  
GVRATKTSERQPRGRQPIKVRQPBKTAQPGYEMPLYLGNQMGWAGWLLSPG  
SRPSWPTDPRKRSRLNKVLDITLTCGADLNGYIPLVAGLPGARVLAHGVCPVLC  
GVNATGNLPGCSFSIPLIALSLCTTPASAEVNAVSGMYHTVLDGNSNLSVEAAD  
MIMHAPGCVPCVRENSSRCWALTPTLAARASVPTTIRHVDLVGAAGFCAMV  
VGDLCGVFLISLPTFSPRHETVODNCISIVPGHGHMAAMMMMSPTALVY  
SOLARIPQAVVYVAGAHGVLALAYISMAKNNAKVLWLLPFGVDGQTRISGTO  
GHTTGLVLSRSSGSSQNLQLININGSHINIKTALNCDSLNTGFLALPYTHKNAS  
GCPARMACSRISDITFAQMGPIYTAQPSISDQPCMYAPQCGIIVASQVCGVLC  
FTSPVAVGTTDRSGVPTVSGENETDVLINTRPQGMNFGCTMANSSTGTTKCG  
PCNIGGVNNLTCTPCDFRKGPEATYKCGSGPMLPRCLVDVRYLMLPCTCNV  
SIFKRYVVGVEHRLNAACNWTBGRGLDRLDSELSPLILSTTEMQNLPCSTTLL  
PALSTGLIHHQNVVDVQIVYIGSAVSVFYKNDYVYLLFILLADAVCAQIMMMLL  
LQAQALLENLVVLAASVAGHGLVLFVFCAMVYIKGRIVPGAALYALGVWVLL  
LLALPRAVYAMREMAASCGGAVLVGMLITLSPYKVFILRLIMLQYILITRAEHL

LOVWPLNVNGBDAILFLTCATHEPILFDITKLLAILGLMVLQAGITRVFVFR  
AAGLIRACMLRVKXGAYVOMAFPKALIFPQYVYHDLTPLRGAHATLVAVE  
PVPSDMEETKTIITMGADPAAGDIIIGLPVARRGRRLILDPALISBOGRILALPIT  
AYSQOTRLDLCIITSLTGRKNOVBGVQVYVSTATQTFLATCYNQVCTYPRHAGSK  
TLAGKGPITQMYTNVDDLVGMQAPBARSMPTCTGSSDLVLTVRADVIIVRRRG  
DGRSLSPRVSVYIKSGSGGLCPGSHVVGIIIRAACTGVAKADVPFVEMMETT  
MRSPFTDNGSPPAVQPTQFVAHLHAPGSSKSTKVPVAAQCYKXVLVNSVYATL  
FGVAMSKAGSDINIRTVARTITTGAPATITSTYKGLFADGCGGADIIICBCHS  
TDSPTIIGITVLVDAEAGARLVIANAPRGSVTVAPNIEVALSNQETIPYK  
ALPIETITGSHLIFGSKKCKDELAKLSLGTINAVAYRGLVSVYIPSGDVVVA  
TDLMTGTGDFDVIDNTCVTQTPVDSPLTPTIDTTPVQDAVSTQRRGRG  
RRGIYRFTVTPBERSGMFDSSVLCBQDAGAMVELPETSRLRALNTPGLPVC  
DHLFEWEGFTGLRIDAHPLSQKQAGDNPPVAYQATCAQAQAPPSMDMKC  
LTRKPLHAGPTPLTYRGAONEVTLTHPRTVTKIMTCSADLVSTVYVGVVLA  
ALAAVCLTGSVIVIGRIISGRALIIPDEVLVIREPMECSOLPYTBQMOLE  
QPKOKALGLTATKQABAAAPVYBSKKRALBAWGMNFISGITYLALSTLPCN  
PALASLMAFTASVTSPLTQSTLLFNILGVAQAQAPASASAFVGAIGAAVGI  
GLGKVLVDILAGYAGVAGALVAFKVMSEMPSTEDLVNLPALISPGALVGVCA  
ILRRHVGEAGVOMNRLLIAPASGRHVSPTHVPESDAARVTOILSNLTQILR  
RLHOMINDECTPCSGSWLRDVMDCITVLSDPFTKLOSLPLPLPGVPLSCORGYK  
GVMRGDGMOTTCPCGQAIQGVHNGSMRIVGPTCSMTWGTTPINAYTTGCTPSP  
APNYSBALMRVAEEVYVTRVGDPHYVTTGCTDNVCKPCVPAPEFTVDGRRLR  
YAPACKPLREEVTFQVGLNOVLVGSOLPCBEPEDVAVLITMLTDPSTITAEYARL  
ARGSPSLASSASQOLASPLKACTTHDSPPADLLEANLIMQEMGNITRVESN  
KVLILDSFDPLRAEEDEREISVPAEILRKPKFPALPTWAPDPYNPLLESMDPDY  
VPPVHGCPLPPTGKPLPPRRKRTVLTSTYSALAEATKTPFGSGSADSGT  
ATAPPDQSDGSDVESYSMWPLEGEGDLDISGMSVTSSEAGDEPVCSWS  
YTTGALITTPCAEBSKLPINALNSILRHNNMYATTSRACOROKYTTDRLOVLD  
DHYDLKEMKAKASTYAKLSVBEACKLTPPHSASKFYGKADVNLSKAVNHI  
RSVKDLLEDDETETINTVMKANEVFCQPEKGRKAPRLVFPDLGRVCEKALYD  
VESIYOCDDLAPEAROVYRSLTERLYVSGPLTNSKQNGCYRARCASVLTSGNTL  
TCYLKASACGAALDQCTMLVNGDILVVICESAGTQEDASLEFTEAMRYSAAPG  
DPEPEVDLEITTCSSSVSYAHNASGRVYVITRDPTEPLARAMEYARTRPVNWL  
GNITMFTPLMARNILMTHFYSIILADQSKALDQYGAQYSIEPLDQIITRHL  
GLSPFSLHSYSPGEINRVASCLRLKVLGPPLVWRHARSVAKLLISQGRATQSKYL  
FNMVARTKLKLPVPAASRLDLSGMPFVAGYSGDIYHLSRBRBRMFWLCLLLSVGV  
GYLLPMR"

ORIGIN  
Query Match 86.1%; Score 891; DB 14; Length 9374;  
Best Local Similarity 91.3%; Pred. No. 2.4e-195;  
Matches 945; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
Qy 1 ACAGAGTGAATGGGGTGGCTGCAGAGAGACGCTCGGGGTGAACCTCTCTACGG 60  
Db 6686 ACTGAGTGAATGAATGGCTGCAGAGTGCCTCGGGGTGAACCTCTCTACGG 6755  
Qy 61 GAGGAGTCAATTCAGAGTGGGCTCAACCAATACCTGTTGGGTGCGACCTCCATGC 120  
Db 6756 GAGGAGTCAATTCAGAGTGGGCTCAACCAATACCTGTTGGGTGCGACCTCCATGC 6815  
Qy 121 GAGCCCGAACCAGATGTAAGAGTGTCACTTCATGCTCACCGAACCTCCCAATCA 180  
Db 6816 GAGCCCGAACCAGATGTAAGAGTGTCACTTCATGCTCACCGAACCTCCCAATCA 6875  
Qy 181 GCAAGAGCGGTAAAGCGAGGCTGGCAAGGAGGCTCCCGCTCTGGGCGAGCTTTCA 240  
Db 6876 GCAAGAGCGGTAAAGCGAGGCTGGCAAGGAGGCTCCCGCTCTGGGCGAGCTTTCA 6935  
Qy 241 GCTAGCCAGTTGTCTGGCCCTCTCTCGAAGCGCATTAATTAACCAATGACTTCCA 300  
Db 6936 GCTAGCCAGTTGTCTGGCCCTCTCTCGAAGCGCATTAATTAACCAATGACTTCCA 6995  
Qy 301 GACGCTGATCTCATGAGGCGCAACCTCTGTGGCGGCGATGATGGGCGGGAACATTAC 360  
Db 6996 GACGCTGATCTCATGAGGCGCAACCTCTGTGGCGGCGATGATGGGCGGGAACATTAC 7055  
Qy 361 CGCGGAGTGAAGGAACAAGTAGTATACTCTGACCTTTTCCAGCCGCTCGAGCGGAG 420  
Db 7056 CGCGGAGTGAAGGAACAAGTAGTATACTCTTGGACTTTTTCGATCGGCTCGGCGGAG 7115  
Qy 421 GAGGATGAGCGGGAAGTGTCCGTCCCGCGGAGATCTCTGCGGAATCCAGAAATTTCCA 480



```

Db      7116 GAGGATGAGAGGAAATATCCGTTCCGCGGAGATCTCTGGGAAACCCAGGAAGTCCCC 7175
Qy      481 CCAGCGATCCCCGATGGGACGCGCGGATTACACCTTCGCTGTGAGTCTGGAG 540
Db      7176 CCAGCGTTCCTTATGGGACGCGCGGATTACACCTTCGCTGTGAGTCTGGAG 7235
Qy      541 GCGCCGATAGCTCCCTCAGTGTACATGGGTGCCACTGCGCACTACTAGACCCCT 600
Db      7236 GACCCGAGTACGCTCCCTCGGTGGTGCAGGCGGTGCCATTGCGACCTACCAAGGCGCT 7295
Qy      601 CCTATACCACTCCAGGAGAAAGAGACAGTGTCTGTGACAGATTCACCGTCTTCT 660
Db      7296 CCATTACCACTCCAGGAGAAAGAGACAGTGTCTGTGACAGATTCACCGTCTTCT 7355
Qy      661 GCGCTGGCGAGAGTTCGCAAAAGCTTTGTAGTCTCGGACCGTGGCGGTGACAG 720
Db      7356 GCGCTGGCGAGAGTTCGCTAAGACCTTTGGGAGCTCCGGGTGCTGGCGCGGACAGT 7415
Qy      721 GGGACGGCAACGCGCCCTCTGACCAATCTCTCGGACGCGCGGAGACCGCGATCTGACGTT 780
Db      7416 GGGACGGCAACGCGCCCTCTGACCAATCTCTCGGACGCGCGGAGACCGCGATCTGACGTT 7475
Qy      781 GAGTGTATTCTCCATGCGCCCTTGAAGGGGAGCGCGGAGACCGCGATCTGACGAGC 840
Db      7476 GAGTGTATTCTCTCCATGCGCCCTCTGAGGAGGAGCGCGGAGACCGCGATCTGACGAGT 7535
Qy      841 GGGTCTGTGCTACCGTGAAGTGAAGAGCGCGGTGAGAGCGTCTGCTGCTGATGATGTC 900
Db      7536 GGGTCTGTGCTACCGTGAAGTGAAGAGCGGTGAGAGCGTCTGCTGCTGATGATGTC 7595
Qy      901 TACACATGACAGGCGCTCTGATACGCCATGCGCTGCGGAGAAAGCAAGCTGCCATC 960
Db      7596 TACACATGACAGGCGCTCTGATACGCCATGCGCTGCGGAGAAAGCAAGCTGCCATC 7655
Qy      961 AACGCGTTAGCAACTCTTGTGCTGCTGACCAACATGATGATGATGATGATGATGATG 1020
Db      7656 AACGCGTTAGCAACTCTTGTGCTGCTGACCAACATGATGATGATGATGATGATG 7715
Qy      1021 AGCGCAAGCCAGCGG 1035
Db      7716 AGCGCAAGCCAGCGG 7730

RESULT 14
AF165059 9379 bp RNA linear VRL 04-NOV-1999
LOCUS Hepatitis C virus strain MD8-1 complete genome.
DEFINITION AF165059
ACCESSION AF165059
VERSION AF165059.1 GI:5918956
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirinae.
REFERENCE 1 (bases 1 to 9379)
AUTHORS Nagayama,K., Kurotsaki,M., Enomoto,N., Maekawa,S.Y., Miyasaka,Y.,
Tazawa,J., Izumi,N., Marumo,F. and Sato,C.
TITLE Time-related changes in full-length hepatitis C virus sequences and
hepatitis activity
JOURNAL Virology 263 (1), 244-253 (1999)
MEDLINE 20013325
PUBMED 10544098
REFERENCE 2 (bases 1 to 9379)
AUTHORS Nagayama,K., Kurotsaki,M., Enomoto,N., Maekawa,S., Miyasaka,Y.,
Sakamoto,N., Fukusaki,T., Tazawa,J., Izumi,N., Marumo,F. and Sato,C.
TITLE Direct Substitution
JOURNAL Submitted (06-JUN-1999) Second Department of Internal Medicine,
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,
Tokyo 113-8519, Japan
FEATURES
source 1..9379
/organism="Hepatitis C virus"

```

```

CDS
/mol_type="genomic RNA"
/strain="MD8-1"
/db_xref="taxon:11103"
/notes="from a patient with persistently high serum ALT
values; patient 8, point 1
genotype: 1b"
330..9362
/codon_start=1
/product="polyprotein"
/protein_id="AAD56194.1"
/translation="MSTNPKPQKRTKRNTRRNPQVDFPGGQIVGVLLPRGRPL
GVATRTKRSRSPQRPQIPKAPRPERTRTAQVPEPVLNGBMGMAWLSRPG
SRPMGNDPRRNRNLGVITDTTCGPRADLMGYLPIVAGPLGAAARLAAHVRLD
SVNVAITGNLPGCSFSIFLLALSLCTIPASAEVKNVSVHTVNDCCNSIITYEAD
ILMHPGMPQVREONIISRCVVALPRLAARI SVPTTIRHVLVLGTAFCSAMY
VDDLGSEVLVQLTFPSRPHETVODCCISYPGVSGIRHMMMDMMMSPTALV
SOLLRIPQAVVDMVAGAHVGLAGLVAVSMGNMAKVLIMLLFAGVDGTVTGTO
ASAHILSPFLTVGPSOKILOVNTGSMHINRTALNCDSILQTFLLALFYTHFNAS
GCPERMASSRPIDTDPQCHGPTTYAEPNNPDRPYCMHAPRPGCIERASKCPYIC
FTPSVAVGTDRFQVPTNMGNEETDVLNNTRPQGNWFGCTMNSTGTCTCGG
PPCNIAGAGNNLLICPTDCFRHNPATYTKCSGWLPRCLVDYPRILMHPCTVNF
TIFKRYMYGVGEHRLSAACMTNTRGRCLEDRDSELSPLISTTEMOILPCSFPTL
PALSTGLHLHRNIVDOVYLGVGAVVSPALIRMEYVLLFLPLADARCAQIMMLL
IAQARAILENIVLNAASVAGAHGILSLFVPCAMAYIOGRIVPGAAVAVGVMPPLL
LLALPPRAYAMDREMASSCGAVVGLAFILSLPYRYRYFLARLIMWLOYPITREAH
LDVWVPLNVRGGRDAILILMCVHRELFDITLKLILFGLMVLQGITVPIVFR
AHGLIRACMLVRKAVAGHVQMAFVKLAALNTGYVLDHAPLQMAHSGRLDAVFA
PVFSQDMETKILITWGDPAACGDIILGLPVSARRREILGLPDSLEGGMFLAPIT
AVSOOTRDLGCIITSLNGRDNGVEGVOVSTNOSPLACINSCGVCTVPHGASK
TLAGKSPITQWTVNTDVLVGMQAPRASMTPCTGSSDLYLVTRHADVIVRRG
DGRGLSBRPVSYIKGSSGRLPLPSGHAACVIFPAACVTRVAAVDFIPRESMETT
MKSVPFTDSSPPAVQVTEVQVHLHAPTSGSTVPPAAYAOGRKVLNLYSPVETL
GGGAVMSKAYGTDPMIRGITRTTGAAPTYSYTKFPLADGCGSGAADIILICDECHS
TDSITLIGITVLDQAEAGRLVVLATPESGTVPPHPIEVALSTGBIPEFGH
APIRVIKGRHLIFCHSKKCDLAAKLSALANAAVAYRGLDVSIVPTSGDVVVA
DALMNGYDGFDSVINDCTVOTVDVSDPTFTIETTVPQDASVRSRGRGRG
RCGITRPTPRGSRPMBSSVLCRCYDAGCAWYLTAEFTSVIRAYINTGDLPPCO
DHLFEWSEFTGLTHIDAFILSOTKQAGNPFYLAQVQVACARQAPPSMDQMLK
LIRLKPFTLHPTPLLYRLGAVQNEVTLHPITKFMACNSADLEVTVSVGVGLA
ALAAVCLTGGSVIVIRIILSGKPAVIRPREVLREPEDEMECASHLPYIBQMOLAE
OFKOKALGLOTATKOAEAAAVESKRALETPAKHMMNFISGIVLAGISTLPGN
PALASIMAFSTASITSLPTTOTLMMNIIGGWAQAOLAPPSASAPFVGIADAAGVIS
GLGKLVLDLAGYGAVALVAFVNGSEMSTEDLILPALISRGALVGVCAAI
ILRHVGRGSAVQMMNRLLAFASRGNHVSPTHYIPESDAKAVQIUSLITLQILK
RLHQMINDCSTPCSGSMRLQDKWMI CVLSPEFKWLOSKIMPLRPGVFLCQKRP
GVMRGDGMQITPCGCAOITGHVKNKSMRIVGPKTCSNTHGTPIINATTPCTSP
APNYSRALMRVAAEVEYTRVGDHYVGTMTDVKCGCOVAPAEFTFVGVRLHR
VAPACKPLIREVTVQVGLNOVLYNGSOLPCREBPVAVVTSMLTDPSPHTATKARLL
ARGSPSLASSASQSLASLAKATTTTHDSADALILEANLIMRQMGNTIRVSEN
KVVILDSFDPLAEEDERBVSAALILRSRKFPPALPVMAADNPULFSGSAVSDPY
VPPVHVGCPLEPTKAPRIPPRKRRTVLTBSTVSALAELAKTFKFGSGSAVSDPY
ATAPDQASDDKSDVESVSMPLREBPDPDLSGDSMTVEEASEDVCCMS
YTMGALITPCAAEESKLPINPLSNLSLHHNMVATTSRSASLQKQVTPRLQVLD
DHYRVLKEMKAKATYVAKLISIEACKLTPRHSRKGFGYADVSISSKAYNI
RSWMDLLEDTETPIDITIMAKSEVFCQPEKSKKPARLIYFPDLGVRCVCKALYD
VSTLPLQAMSSGYFOYSPKQVPEPLVNTWKSCKPMPGFSYDTCFSTVYESIIRT
TESIYOCCLDAPEARKAKISLTERLYIGGPLNNSGOMCGRACASGLTTSNCNTL
TCYLKASACRAKAKODCTMLVNGDVLVIVICSSAGTQDASLARTFEMTYSAPG
DPPOPEYDELITSCSNVSVAHNDGSRVYVLTDPPTPLARAMEBTRHPNVSMW
GNIMVAPLMAHMLTMHPISILIAQOLEKALDCQIYGACYSIEMPTDIPOTIRRH
GLSARSLHYSFPEILNRVASCURKLGVPFLRWRRHARSVRKALLSGGRATCGKYL
FNMAVRTKLKLTPIPAASRLDLSGFVAGVSGDLYHLSLRARPRPMFWMLLSVGV
GIYLLPNR"

```

## ORIGIN

```

Query Match      86.1%; Score 891; DB 14; Length 9379;
Best Local Similarity 91.3%; Pred. No. 2.4e-195;
Matches 945; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
Qy      1 ACAGAAAGTGAAGGAGTGGGCTGCAAGAGTACGCTCCGCGCTGCAAACTCTCTCAACG 60
Db      6696 ACAGAAAGTGAAGTGAAGTGGCAAGTACGCTCCGCGCTGCAAACTCTCTCAACG 6755

```

61 GAGAGGTGCATTCCAGGTTCGGGCTCAACCAATACCTGTTGGTCCAGCTCCCATC 120  
6756 GAGGAGGTGCATTCCAGGTTCGGGCTCAACCAATACCTGTTGGTCCAGCTCCCATC 6815  
121 GAGCCGGAACCGGATTTAGAGTGTCTACTTCCATGCTCAACGACCTCCCAATCACA 180  
6816 GAGCCGGAACCGGATTTAGAGTGTCTACTTCCATGCTCAACGACCTCCCAATCACA 6875  
181 GAGAGAGCGGCTAAGGCGAGGCTGCGCAGGCGGCTCTCCCTCCCTGGCCAGCTCTTCA 240  
6876 GAGAGAGCGGCTAAGGCTGCGCAGGCGGCTCTCCCTCCCTGGCCAGCTCTTCA 6935  
241 GCTAGCAGTTGTCTGGCCCTTCTCTGAGAGCGACATACATTAACTTCTCCCA 300  
6936 GCTAGCAGTTGTCTGGCCCTTCTCTGAGAGCGACATACATATATATCTCCCG 6995  
301 GACGCTGACCTCATCGAGGCTCAACCTCTGTGCGGCGATGAGATGGGCGGGAATTACC 360  
6996 GACGCTGACCTCATCGAGGCTCAACCTCTGTGCGGCGATGAGATGGGCGGGAATTACC 7055  
361 CGGTGGAGTCAAGAGCAAGGATGATCTGGAATCTTTGACCCCGCTCGAGCGGAG 420  
7056 CGGTGGAGTCAAGAGCAAGGATGATCTGGAATCTTTGACCCCGCTCGAGCGGAG 7115  
421 GAGGATGAGCGGGAAGTGTCCGTCCTCGGCGGAGATCTCGGAAATCCAAAGATTTCCA 480  
7116 GAGGATGAGAGGAGGATGATCTCGTGGCGGGAATCTCGGAGATCCAGAAAGTTCCCC 7175  
481 CCAAGCATCCCGCATGGGCGAGCCCGGATTTACAACTTCGCTGTGAGATCTCGAAG 540  
7176 CCAAGCATCCCGCATGGGCGAGCCCGGATTTACAACTTCGCTGTGAGATCTCGAAG 7235  
541 GCCCGGAGTACGTCTCCAGTGTGATCGGATGATCGGATGATCGGATGATCGGATGAT 600  
7236 GACCCGAGTACGTCTCCAGTGTGATCGGATGATCGGATGATCGGATGATCGGATGAT 7295  
601 CCTATACCACTTCAACGAGAAAGAGAGATGTTCTGACAGAAATCCACGCTGCTTCT 660  
7296 CCAATACCACTTCAACGAGAAAGAGAGATGTTCTGACAGAAATCCACGCTGCTTCT 7355  
661 GCCCTGGCGAGAGTTGCCAAGAGCTTTTGGTGAAGTCTCGGACCGTGGCGCTCGACAG 720  
7356 GCCCTGGCGAGAGTTGCCAAGAGCTTTTGGTGAAGTCTCGGACCGTGGCGCTCGACAG 7415  
721 GGCAGCGCAACCGCCCTCTGACCAATGCTCCGACGAGCGGCGAGCAAGATCTACGTT 780  
7416 GGCAGCGCAACCGCCCTCTGACCAATGCTCCGACGAGCGGCGAGCAAGATCTACGTT 7475  
781 GAGTGTATCTCTCATGCTGCCCCCTTGAAGGCGAGCCGCGGAGACCCGATCTCAGCGAC 840  
7476 GAGTGTATCTCTCATGCTGCCCCCTTGAAGGCGAGCCGCGGAGACCCGATCTCAGCGAC 7535  
841 GGGTCTTGTCTCTCATGCTGCCCCCTTGAAGGCGAGCCGCGGAGACCCGATCTCAGCGAC 900  
7536 GGGTCTTGTCTCTCATGCTGCCCCCTTGAAGGCGAGCCGCGGAGACCCGATCTCAGCGAC 7595  
901 TACACATGACAGGCGCTGATGATCGCATGCGCTGCGGAGGAAAGCAAGTGGCCATC 960  
7596 TACACATGACAGGCGCTGATGATCGCATGCGCTGCGGAGGAGAGCAAGTGGCCATC 7655  
961 AACGGCTTGAAGCACTTTTGTGCTGATCAACCAATGATGCTAGGCTTACCAATCCGCG 1020  
7656 AACGGCTTGAAGCACTTTTGTGCTGATCAACCAATGATGCTAGGCTTACCAATCCGCG 7715  
1021 AGCGCAAGCGACCGG 1035  
7716 AGCGCAAGCGCTGCGG 7730

RESULT 15  
API65060  
LOCUS AF165060 9379 bp RNA linear VRL 04-NOV-1999

DEFINITION Hepatitis C virus strain MD8-2 complete genome.  
ACCESSION AF165060  
VERSION 1  
KEYWORDS GI:5918958  
SOURCE  
ORGANISM Hepatitis C virus  
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
REFERENCE 1 (bases 1 to 9379)  
AUTHORS Nagayama, K., Kurosaki, M., Enomoto, N., Maekawa, S. Y., Miyasaka, Y.,  
Tazawa, J., Izumi, N., Marumo, F. and Sato, C.  
TITLE Time-related changes in full-length hepatitis C virus sequences and  
hepatitis activity  
JOURNAL Virology 263 (1), 244-253 (1999)  
MEDLINE 20013325  
PUBMED 1054098  
REFERENCES 2 (bases 1 to 9379)  
AUTHORS Nagayama, K., Kurosaki, M., Enomoto, N., Maekawa, S., Miyasaka, Y.,  
Sakamoto, N., Fukuma, T., Tazawa, J., Izumi, N., Marumo, F. and Sato, C.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-1999) Second Department of Internal Medicine,  
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,  
Tokyo 113-8519, Japan  
FEATURES  
source  
location/Qualifiers  
1..9379  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/strain="MD8-2"  
/db\_xref="taxon:11103"  
/note="From a patient with persistently high serum ALT  
values; patient 8, point 2  
genotype: 1b"  
330..9362  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="A056195.1"  
/db\_xref="gi:5918958"  
/translation="MSTNPKPORKTKNTNRPODVFPGGQIVGGVLLPRRGLR  
GVATRKTSERSOPRRGROPKARREGRWMAQGPWMPVKGSGMGAGMLSPG  
SRPWSGNDPRRSRNIGKVIIDTLCPADIMGYIPIGVGAPIGAAALAHGVAVLD  
SVNYATGNLPECSFSTFLALISGLTTPASATYERANVSQYHTVNTDSSNSITVEAD  
IIMHTPGCTPCVRBGNISRCVVALPTLALANISVPTTTLRRHDLVGTAIFCSAY  
VGDLGSEFVLSQLFPSPRHETVODNCISYGHSGHGMADMMNMSPTALVY  
SOLLRIPQAVVAVAGAHMVGALAYSCNMGNAKVLIMWLPAFGVGTTHVVGAG  
AHSATMSPTTVGSPXKIDLVNTNGSHMIRTKLNCNDSLQTFGLAALFTTHFNS  
GCEPRMASCRIPIPDQGWPIITYAEKNSDPRQRCMYAPRPGISIPASVGVPC  
FTSPVYVGTTRDFGVPTTKMGEKNTVLLNTRPPQGNFECTMAWSTGFTTCG  
PPCNIGAGANTLTCTPDCERKHPDEATYTKCGSPWLPRLCLVDYPRLMYPCVNP  
TIIFVRYVVGVEHRLSAAQWTRGECDELDROBSLSPLLSTTMOQLPCEFTL  
PALSTGLIHLEHRIIVDOYLYGVGSAAVSIIRWEYVLLFLDLADARVACLMMLL  
IAQAEALLENLVNNAASVAGAGILSFVFCAMYIKGRLVGAAYALGVMPLL  
LLALPRAYVMDENMAASCGGAFVGLAFLTSPYKVLPLRLIMLYEFTPEABH  
LQWVPPPLNVRGDAIILMCAVHPHELIDITLLALIREPLMVLQAGITRVYPR  
AHGILRACMLVRKAVGAYVOMAFVKLAALGYVYVHLAFLQHMASGLDLVAYR  
PVPSSDMETKXITWGAADTAACGDIILGLPVASARGRLILGLPADLSLGGQRLAPIT  
AYSOOTRGLGLCITSLTGRDKNOEVEVQVSPATOSPLATCNGCMTVFRGAGK  
TLAPKBPITQMYTNPODLYGMOAPGASMTCTCGSSDLYLVTHADVIPIRAG  
DGRSLISPRVSVLKGSSGGLCPGHAAGVIGRAVCRGVAKAADPIPEVEMET  
MRSPVPTDNSPPRAVPTPOTROYAHHAATPGSGSKRVPAAYAAQGYKVLVNPSTAL  
GFGVWMSKAYGTDPNIRGTITTTGAPITYSTGKFLADGSGGAYDITICBGS  
TDSYTLIGITVLDQAEYTAGARLVLTATPPGSAVPHPIEBVALSNGEIPYK  
AIPIEVKGGHLLFCHSKKKCDLALAKSLGLNAAAYKGLDVSYPISGDAVVA  
TDLAMTGYTDFSDVINCNTQVOTVEFSDPTTITTTTVPDQAVRSORGRGTG  
RGGIYRPTVBERSPGMPDSVLCEDYDAGCAWELPAETSVSLRAYLNPGLPVO  
DHLEFMSVFTGLTHIDAHPLSOTKQAGNFPYVAYQATYCAAQAPRPSNDWMC  
LIRLPTLHGFTPLVRLGAVONVLTLLPITKTKIMCWAADLHVSTWLVGVVLA  
ALAAVCLTGSVVIVGRIILSGKPAVLPDRVLVQEDMEBCASHLYEQGQVLA  
QFKOKALGLIQTATKQAEAPVVESTRALETFMAMNMFISGIQYLAGISTLPGN  
PAISLMAFTASITSPLTOTTLTLENILIGVAAQLAPSPASAFVAGIGAVGSI  
GLGKVLVILIAGVAGVAGALVAFKWSGEMSPREDIINLPAIISGALVGVCA  
ILRRHVGBGAVVMMNRLLIAPARKNHVSPTHVPPSDAAARTQLISLTITIQULK  
RLHOWINBDCSTPCSSWMLNDWMTCTVLDFETWLOSKIMPLPQVPLSCRGTK  
GWRGDDIMOTTCPGAQITGHVKNRMRIVGPTCSNTWHTGTPINAYTTGPTCPSP



ANVSHALNRVAEEYEVTRVGDFFHVTGTTDNVNCPCQVPAPEFFTEVDVRLHR  
 YAPACKLREBEVTFQVGLNOYLVGSGLOPEBEPDVAVLITSMITDSHITAEKARL  
 AAGSPRLSSSSSSQSLASBIRKATCTTHHDSPPADLIENALIRORQEMGNTVBSEN  
 KAVILDSFPLRAEBDEBREVSAEILRSRKPPLPVPWARPPLPLESKMDPY  
 VPPVVAAGCELPPTKAPPIPPRRKRTVLTSSSVSALAEIARITGSSSSVSDGT  
 ATAPPDGSDGDKESDVSYSMPLEGGPDLSDGSMATVSEASDVCCT  
 YMTWGLITPCAEESKLPINPLNSLRHNNVATSTSAELRQKTVFDLOVLD  
 DHRDVLKEMKAKASTVAKLISIEBCKLTPHSAKSKRGVAKDVRSLSSAVNHI  
 RSVMBLLEDETERPIDTTIMAKSEFVCPYBKGRRPARLIVPDIGVRCECMALYD  
 VSTLPOAVWSSYGFQYSPKQREFLVWTKSKCPMGSYDTRGCDSTVTSIDRT  
 EBSIYOCCLAPARAKIISLTERLYIGGLTNSKQNCGRCRASGVLTTSIGMTL  
 TCYKATAACRAKLODCTMLVNGDDLWVYCESAGTOEDASLRVFEAMRTVSAAPG  
 DPEQPEYDELITSCSNVSVADHDSGRVYVTRDPTPLAAMETABRTVNSML  
 GNTIWAYPTLMAMIIIMTFSLIAOEOLKALDOIVACGSIPIBLDPOIIEBLH  
 GLSAFSLHSYBGEINRVASCLRLKGVPIPRVRRHARSVRATLISQGRAPATCGKTL  
 FMAVATKLTLPITPAASRLDLSGMFVAGISGDIYHLSIRARPRMFMCLLISVGV  
 GYILPNR"

## ORIGIN

Query Match 85.9%; Score 889.4; DB 14; Length 9379;  
 Best Local Similarity 91.2%; Pred. No. 5.6e-195;  
 Matches 944; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ACAGAGTGATGGGGTGCAGCTGCAAGGTAGCTCCGGCGTGCACAACTCTCTTACGG 60  
 DB 6696 ACAGAGTGATGGAGTACGCTGCAAGGTAGCTCCGGCGTGCACAACTCTCTTACGG 6755  
 QY 61 GAGAGGTACATTCAGGTGCGGCTCAACCAATACCTGGTGGGTGCGAGCTCCATGC 120  
 DB 6756 GAGAGGTACATTCAGGTGCGGCTCAACCAATACCTGGTGGGTGCGAGCTCCATGC 6815  
 QY 121 GAGCCGGAACCGGATGAGAGTGTCACTTCCATGCTACCGACCCCTCCCATGACA 180  
 DB 6816 GAGCCGGAACCGGATGAGAGTGTCACTTCCATGCTACCGACCCCTCCCATGACA 6875  
 QY 181 GCAGAGACGGCTAAGCGCAGCTGCGCAGGGGGTCTCCCTCTCTGGCAGCTCTTCA 240  
 DB 6876 GCAGAGACGGCTAAGCGTACGCTGCGCAGGGGGTCTCCCTCTCTGGCAGCTCTTCA 6935  
 QY 241 GCTAGCCAGTGTCTGGCCCTTCTCTGAGGCGACATACATTAACCAAAATGACTTCCA 300  
 DB 6936 GCTAGCCAGTGTCTGGCCCTTCTCTGAGGCGACATACATTAACCAAAATGACTTCCA 6995  
 QY 301 GAGGCTGACTCATTCAGGCGCAACTCTCTGGCGGATGAGATGGCGGGGACATTAAC 360  
 DB 6996 GAGGCTGACTCATTCAGGCGCAACTCTCTGGCGGATGAGATGGCGGGGACATTAAC 7055  
 QY 361 CGGCTGAGTCAAGAAACAAGTAGTAACTCTGGAATCTTTCGACCCGCTCCGAGCGGAG 420  
 DB 7056 CGGCTGAGTCAAGAAACAAGTAGTAACTTGGAACTCTTTCGACCCGCTTTCGAGCGGAG 7115  
 QY 421 GAGGATGAGCGGGAAGTGTCCGTCGCGCGGAGATCTGCGGAAATCCAGAAATTTCCA 480  
 DB 7116 GAGGATGAGCGGGAAGTGTCCGTCGCGCGGAGATCTGCGGAAATCCAGAAATTTCCA 7175  
 QY 481 CCAGCGATGCCCGCATGGGCAAGCCCGGATTAACAACCTCTCCGCTGCTGAGTCTGGAAG 540  
 DB 7176 CCAGCGATGCCCGCATGGGCAAGCCCGGATTAACAACCTCTCCGCTGCTGAGTCTGGAAG 7235  
 QY 541 GCCCGGGAATCACTCCCTCCAGTGTATCATGGGTGCGCACTGCGCACTTAAGAAGCCCT 600  
 DB 7236 GCCCGGGAATCACTCCCTCCAGTGTATCATGGGTGCGCACTGCGCACTTAAGAAGCCCT 7295  
 QY 601 CCTATACCAACTTCACGGAGAAAGAGGAGTGTCTGA CAGAAATCCAGCTGTCTTCT 660  
 DB 7296 CCTATACCAACTTCACGGAGAAAGAGGAGTGTCTGA CAGAGTCCAGTGTCTTCT 7355  
 QY 661 GCCCTGGCGGAGCTTGGCAAAAGCTTTTGGTAGCTCCGGAACCTGCGGCGCTGAGACG 720  
 DB 7356 GCCCTGGCGGAGCTTGGCAAAAGCTTTTGGTAGCTCCGGAACCTGCGGCGCTGAGACG 7415  
 QY 721 GGCACGGCAACCGCCCTCTGACCAATCTCCGACGACGCGGAGCAGAGATCTGACGTT 780

DB 7416 GGCACGGCACTGCTCCCTCCGACAGGCTCCGACGAGCGGCAAGAAATCCAGCTT 7475  
 QY 781 GACTCGTATTCCTCCATGCCCCCTTGAAGGGGAGCCGGGGGACCCCGATCTCAGCGAC 840  
 DB 7476 GACTCGTATTCCTCCATGCCCCCTTGAAGGGGAGCCGGGGGACCCCGATCTCAGCGAC 7535  
 QY 841 GGGCTTTGGTCTACCGCTGAGTGAAGAGCGCGTGAAGAGCTGCTGCTGCTGATGTC 900  
 DB 7536 GGGCTTTGGTCTACCGCTGAGTGAAGAGCGCGTGAAGAGCTGCTGCTGCTGATGTC 7595  
 QY 901 TACACATGACAGCGGCTTGATCAGCGCAATGCTGCGGAGAAAGCAAGTGGCCATC 960  
 DB 7596 TACACATGACAGCGGCTTGATCAGCGCAATGCTGCGGAGAAAGCAAGTGGCCATC 7655  
 QY 961 AAGCGTTGAGCAACTTTGCTGCTGCTCAACCAATATGTTACAGCTACCATCCCG 1020  
 DB 7656 AAGCGTTGAGCAACTTTGCTGCTGCTCAACCAATATGTTACAGCTACCATCCCG 7715  
 QY 1021 AGCGCAAGCGGCGG 1035  
 DB 7716 AGCGCAAGCGGCGG 7730

Search completed: February 19, 2005, 07:08:43  
 Job time : 4603.2 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 20:55:25 ; Search time 579.412 Seconds  
(without alignments)  
10574.404 Million cell updates/sec

Title: US-09-664-363-4

Perfect score: 1035  
Sequence: 1 ACAGAGATGATGCGGTGGG.....CCCGAGCCGAGCAGCCG 1035

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1035	100.0	2790	2	AAQ46191	Aa46191 PT-NANBH
2	1035	100.0	3372	2	AAQ46202	Aa46202 PT-NANBH
3	1033.4	99.8	1035	2	AAQ12237	Aa12237 Clone UG3
4	1033.4	99.8	1788	2	AAQ46193	Aa46193 PT-NANBH
5	1025.4	99.1	3750	2	AAQ12241	Aa12241 Encodes P
6	1025.4	99.1	7065	2	AAQ46195	Aa46195 PT-NANBH
7	889.4	85.9	7989	10	ADD93733	Ad93733 Hepatitis
8	885.8	85.6	8642	6	ABK8574	Abk8574 Hepatitis
9	884.6	85.5	7987	6	AAD25321	Aad25321 Hepatitis
10	884.6	85.5	7989	6	AAA98968	Aa98968 Hepatitis
11	884.6	85.5	7989	6	AAD25322	Aad25322 Hepatitis
12	884.6	85.5	7989	12	ADJ57845	Adj57845 HCV repli
13	884.6	85.5	7991	6	AA147279	Aa147279 Hepatitis
14	884.6	85.5	7992	6	AA147276	Aa147276 Hepatitis
15	884.6	85.5	7992	6	AA147280	Aa147280 Hepatitis
16	884.6	85.5	8001	3	AAA98967	Aa98967 Hepatitis
17	884.6	85.5	8637	3	AAA98966	Aa98966 Hepatitis
18	884.6	85.5	8638	6	ABK8595	Abk8595 Hepatitis
19	884.6	85.5	8638	6	ABK8577	Abk8577 Hepatitis
20	884.6	85.5	8639	6	ABK8573	Abk8573 Hepatitis

21	884.6	85.5	8649	3	AAA98969	Aa98969 Hepatitis
22	884.6	85.5	9605	6	ABK91424	Abk91424 Hepatitis
23	884.6	85.5	9605	6	ABK91411	Abk91411 Hepatitis
24	884.6	85.5	9605	6	ABK91428	Abk91428 Hepatitis
25	884.6	85.5	9605	6	ABK91425	Abk91425 Hepatitis
26	884.6	85.5	9605	6	ABK91426	Abk91426 Hepatitis
27	884.6	85.5	9608	6	ABK91427	Abk91427 Hepatitis
28	884.6	85.5	10690	6	ABK91243	Abk91243 Hepatitis
29	884.6	85.5	10690	6	ABK91242	Abk91242 Hepatitis
30	884.6	85.5	10690	6	ABK91412	Abk91412 Hepatitis
31	884.6	85.5	10690	9	ACA61697	Ac61697 Hepatitis
32	884.6	85.5	10690	10	ADC83762	Adc83762 PHCNeo17
33	884.6	85.5	10691	6	ABK91423	Abk91423 Hepatitis
34	884.6	85.5	10693	6	ABK91443	Abk91443 Hepatitis
35	884.6	85.5	10693	6	ABK91438	Abk91438 Hepatitis
36	884.6	85.5	11076	3	AAA98965	Aa98965 Hepatitis
37	884.6	85.5	11313	12	ADP86273	Adp86273 Hepatitis
38	884.6	85.5	11313	12	ADP86264	Adp86264 Hepatitis
39	884.6	85.5	11313	12	ADP86271	Adp86271 Hepatitis
40	884.6	85.5	11313	12	ADP86275	Adp86275 Hepatitis
41	883	85.3	7141	6	AAD25333	Aad25333 Hepatitis
42	883	85.3	7789	6	AAD25330	Aad25330 Hepatitis
43	883	85.3	7980	10	ADD93724	Add93724 Hepatitis
44	883	85.3	7980	10	ADD93725	Add93725 Hepatitis
45	883	85.3	7987	6	AAD25329	Aad25329 Hepatitis

## ALIGNMENTS

RESULT 1						
ID	AAQ46191	standard; DNA; 2790 BP.				
AC	AAQ46191;					
XX						
DT	27-AUG-2003	(revised)				
DT	25-MAR-2003	(revised)				
DT	24-FEB-1994	(first entry)				
XX						
DE	PT-NANBH virus BHC-11 fusion protein.					
XX						
KM	Parenterally transmitted non A non B hepatitis; PT-NANBH;					
KM	hepatitis C virus; HCV; NS5; E1; linker; detection; diagnosis; antigen;					
KW	vaccine; BHC-11; replicase; core protein;					
KW	Antograpa californica nuclear polyhedrosis virus; AcNPV; polyhedrin; ss.					
XX						
OS	Hepatitis virus.					
XX						
FH	Key	Location/Qualifiers				
FT	CDS	1..2790				
FT		/*tag= a				
FT		/product= "fusion_protein"				
FT		1..63				
FT		/*tag= b				
FT		/label= AcNPV_polyhedrin_N-terminal				
FT		64..1851				
FT		/*tag= c				
FT		/label= PT-NANBH_NS5				
FT		1852..1875				
FT		/*tag= d				
FT		/note= "Synthetic linker region"				
FT		1876..2706				
FT		/*tag= e				
FT		/label= PT-NANBH_core_and_E1_regions				
FT		2707..2790				
FT		/*tag= f				
FT		/note= "Polyhedrin gene sequence read out-of-frame"				
PN	WO9317110-A2.					
XX						
PD	02-SBP-1993.					
XX						

PF 19-FEB-1993; 93WO-GB000345.  
 XX 21-FEB-1992; 92GB-00003803.  
 XX (WELL ) WELLCOME FOUND LTD.  
 PA Parker D, Rodgers BC;  
 XX WPI; 1993-288415/36.  
 DR P-PSDB; AAR41431.  
 XX  
 PT New recombinant polypeptide for diagnosing hepatitis C - contains three  
 PT distinct antigens from different viral regions, also useful in protective  
 PT vaccinee.  
 PS  
 XX Disclosure: Page 20-24; 99pp; English.  
 XX  
 CC GB-A-2239245, discloses a recombinant polypeptide BHC-11 which comprises  
 CC an antigen obtained from the non-structural coding region (NS) (the 3'  
 CC end) and one antigen from the structural coding region (S) (the 5' end)  
 CC of the NANBH virus. Specifically BHC-11 (AAQ46191) contains a portion of  
 CC the non-structural region of the virus, called NS5, (putative replicase)  
 CC at the N-terminus joined via a synthetic linker to a portion of the  
 CC structural region which contains almost all the core protein sequence (9  
 CC amino acids from the N-terminal are not present) and a part of a sequence  
 CC from the structural region called E1. It is disclosed that BHC-11 may be  
 CC used in diagnosis of PT-NANBH. If at least three different PT-NANBH  
 CC antigens are used to screen for PT-NANBH, the screening is much more  
 CC sensitive as compared to the use of only two PT-NANBH antigens. Pref.  
 CC antigens are described in AAQ46192-94. Two new antigenic regions of the  
 CC PT-NANBH genome are given in AAQ46198-99. AAQ46202 describes an improved  
 CC PT-NANBH recombinant polypeptide. (Updated on 25-MAR-2003 to correct OS  
 CC field.) (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 2790 BP; 585 A; 861 C; 789 G; 555 T; 0 U; 0 Other;

Query Match 100.0%; Score 1035; DB 2; Length 2790;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-268;  
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGAGATGATGGGCTGGCTGCAAGTACGCTCCGGCTGCAACCTCTCTACGG 60  
 DB 64 AAGAGATGATGGGCTGGCTGCAAGTACGCTCCGGCTGCAACCTCTCTACGG 123  
 QY 61 GAGGAGTACATTCAGGTCGGGCTCAACCAATACCTGTTGGGTCGAGCTCCATGC 120  
 DB 124 GAGGAGTACATTCAGGTCGGGCTCAACCAATACCTGTTGGGTCGAGCTCCATGC 183  
 QY 121 GAGCCCGAACCAGATGTAGCACTGCTCACTTCATGCTACCGACCCCTCCCATCA 180  
 DB 184 GAGCCCGAACCAGATGTAGCACTGCTCACTTCATGCTACCGACCCCTCCCATCA 243  
 QY 181 GAGGAGTACATTCAGGTCGGGCTCAACCAATACCTGTTGGGTCGAGCTCCATGC 240  
 DB 244 GAGGAGTACATTCAGGTCGGGCTCAACCAATACCTGTTGGGTCGAGCTCCATGC 303  
 QY 241 GGTACCGATGTCGCTGCTCTCGAAGGCAATACATTAACCAAAATGATTCCTCA 300  
 DB 304 GGTACCGATGTCGCTGCTCTCGAAGGCAATACATTAACCAAAATGATTCCTCA 363  
 QY 301 GAGCGTGAATCATGAGGCAACCTCTGTGGCGCATGAGATGGCGGAGCATTAAC 360  
 DB 364 GAGCGTGAATCATGAGGCAACCTCTGTGGCGCATGAGATGGCGGAGCATTAAC 423  
 QY 361 GCGGTGAGTCAAGAAACAAGTATTAATCTTGTGACCCGCTCCGAGCGAG 420  
 DB 424 GCGGTGAGTCAAGAAACAAGTATTAATCTTGTGACCCGCTCCGAGCGAG 483  
 QY 421 GAGGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTCGGAAATTCAGAAATTCCTCA 480  
 DB 484 GAGGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTCGGAAATTCAGAAATTCCTCA 543  
 QY 481 CCAAGCATGCCCATGAGGACGCCCGGATTAACAACCTCTGCTGAGATCTGGAAG 540

DB 544 CCAAGCATGCCCATGAGGACGCCCGGATTAACAACCTCTGCTGAGTCTGGAAG 603  
 QY 541 GCCCGGATCAATGTCCTCCAGTGTATAGGTCGCCCACTGCACTTAAGACCCCT 600  
 DB 604 GCCCGGATCAATGTCCTCCAGTGTATAGGTCGCCCACTTAAGACCCCT 663  
 QY 601 CCTATACCACTCCAGGAGAAAGAGACAGTGTCTTGAAGAAATTCACCGTCTCT 660  
 DB 664 CCTATACCACTCCAGGAGAAAGAGACAGTGTCTTGAAGAAATTCACCGTCTCT 723  
 QY 661 GCCCTGGGAGCTTTGCCAACAAGCTTTGTGATCTCCGACCGTGGCGCTGACAC 720  
 DB 724 GCCCTGGGAGCTTTGCCAACAAGCTTTGTGATCTCCGACCGTGGCGCTGACAC 783  
 QY 721 GGCACGGACACCGCCCTCCAGCAATCTCCAGCGAGCGAGAGATGTACGTT 843  
 DB 784 GGCACGGACACCGCCCTCCAGCAATCTCCAGCGAGCGAGAGATGTACGTT 843  
 QY 781 GAGTGTATTCCTCATGCCCCCTTGAAGGAGCGGAGGACCCGATCTCAGCAG 840  
 DB 844 GAGTGTATTCCTCATGCCCCCTTGAAGGAGCGGAGGACCCGATCTCAGCAG 903  
 QY 841 GAGTGTATTCCTCATGCCCCCTTGAAGGAGCGGAGGACCCGATCTCAGCAG 900  
 DB 904 GAGTGTATTCCTCATGCCCCCTTGAAGGAGCGGAGGACCCGATCTCAGCAG 963  
 QY 901 TACACATGAGAGCGCTCTATACAGCAAGCCTCGGAGAGAAAGCACTCCCATC 960  
 DB 964 TACACATGAGAGCGCTCTATACAGCAAGCCTCGGAGAGAAAGCACTCCCATC 1023  
 QY 961 AACGGTTGAGCAACTCTTTGCTGCTGACCAACAATGTTCTAGCTACCAATCCCG 1020  
 DB 1024 AACGGTTGAGCAACTCTTTGCTGCTGACCAACAATGTTCTAGCTACCAATCCCG 1083  
 QY 1021 AGCGCAAGCCAGCGG 1035  
 DB 1084 AGCGCAAGCCAGCGG 1098

RESULT 2  
 AAQ46202  
 ID AAQ46202 standard; cDNA to mRNA; 3372 BP.  
 XX  
 AC AAQ46202;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 24-FEB-1994 (first entry)  
 XX  
 DE PT-NANBH NS-NS3-core recombinant polypeptide.  
 XX  
 KM Parenterally transmitted non A non B hepatitis; PT-NANBH;  
 KM hepatitis C virus; HCV; NS3; core; E1; linker; detection; diagnosis;  
 KM antigen; vaccine; BHC-28; replicase; core protein;  
 KM Autographa californica nuclear polyhedrosis virus; AcNPV; polyhedrin; ss.  
 OS Hepatitis virus.  
 XX  
 FH Key  
 FH CDS 1. Location/Qualifiers  
 FT 1..3372  
 FT /tag= a  
 FT /product= "fusion\_protein"  
 FT misc\_feature 1..63  
 FT /tag= b  
 FT /label= AcNPV\_polyhedrin\_N-terminal  
 FT 64..1852  
 FT /tag= c  
 FT /label= PT-NANBH\_NS3  
 FT 1853..1858  
 FT /tag= d  
 FT /note= "Peptide restriction site"  
 FT 1859..2434  
 FT misc\_feature

```

FT      /*tag= e
FT      /label= PT-NANBH_NS3
FT      misc_feature
FT      2435..2457
FT      /*tag= f
FT      /note= "Pari restriction site"
FT      misc_feature
FT      2458..3288
FT      /*tag= g
FT      /label= PT-NANBH_core_and_E1_regions
FT      misc_feature
FT      3289..3372
FT      /*tag= h
FT      /note= "Polyhedrin gene sequence read out-of-frame"
PN      MO9317110-A2.
PD      02-SEP-1993.
XX      19-FEB-1993; 93WO-GB000345.
XX      21-FEB-1992; 92GB-00003803.
XX      (WELL ) WELLCOME FOUND LTD.
XX      Parker D, Rodgers BC;
XX      WPI: 1993-288415/36.
XX      P-PSDB; AAR41439.
XX      New recombinant polypeptide for diagnosing hepatitis C - contains three
XX      distinct antigens from different viral regions, also useful in protective
XX      vaccine.
XX      Claim 7; Page 85-90; 99pp; English.
XX      The NS3 specific region of pDX200 was amplified by PCR using primers D360
XX      and D361 (AAQ46201 and AAQ54882). The obtained fragment was then cloned
XX      into pDX136. Transformants were analysed to identify those which
XX      contained the NS3 sequence inserted in the correct orientation between
XX      the NS5 and core parts of pDX136, this was called pDX208 (AAQ46202).
XX      Recombinant baculovirus BHC-28 was produced. Insect cells infected with
XX      BHC-28 produce antigen NS5-NS3-core. If at least three different PT-NANBH
XX      antigens are used to screen for PT-NANBH, the screening is much more
XX      sensitive as compared to the use of only two PT-NANBH antigens. Pref.
XX      CC antigens are described in AAQ46192-94. Two new antigenic regions of the
XX      PT-NANBH genome are given in AAQ46198-99. AAQ46202 describes an improved
XX      PT-NANBH recombinant polypeptide. (Updated on 25-MAR-2003 to correct PN
XX      field.) (Updated on 27-AUG-2003 to correct OS field.)
SQ      Sequence 3372 BP; 714 A; 1026 C; 954 G; 678 T; 0 U; 0 Other;
Query Match      100.0%; Score 1035; DB 2; Length 3372;
Best Local Similarity 100.0%; Pred. No. 2,8e-268;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ACAGAGTGGATGGGTCGTCACAGGTACGCTCCGCGTGCACAACTCTCTTACCG 60
DB      64 ACGAAGTGGATGGGTCGTCACAGGTACGCTCCGCGTGCACAACTCTCTTACCG 123
QY      61 GAGGAGTCAATTCAGGTCGGGCTCAACCAATACCTGTTGGTTCGACGCTCCATGC 120
DB      124 GAGGAGTCAATTCAGGTCGGGCTCAACCAATACCTGTTGGTTCGACGCTCCATGC 183
QY      121 GAGCCCGAACCAGATAGTCAGTCATCTTCATGCTACCGACCCCTCCACATCA 180
DB      184 GAGCCCGAACCAGATAGTCAGTCATCTTCATGCTACCGACCCCTCCACATCA 243
QY      181 GCAGAGACGGCTAAGCGCAGGTCGCGAGGGGCTCTCCCTCTCTTGCGCAGCTTCA 240
DB      244 GCAGAGACGGCTAAGCGCAGGTCGCGAGGGGCTCTCCCTCTCTTGCGCAGCTTCA 303
QY      241 GCTAGCAAGTTGTCTGGCCCTTCTCGAAGCGACATATACCAAAATGAATTCCCA 300
DB      304 GCTAGCAAGTTGTCTGGCCCTTCTCGAAGCGACATATACCAAAATGAATTCCCA 363

```

```

QY      301 GACGCTGACCTCATCGAGGCAACACCTCTGTGGCGGCAAGAGATGGGCGGAGCATTTACC 360
DB      364 GACGCTGACCTCATCGAGGCAACACCTCTGTGGCGGCAAGAGATGGGCGGAGCATTTACC 423
QY      361 CGCGTGAAGTCAGAGAACAAAGGTAGTAATCTTGACCTCTTTCGACCCGCTCCGAGCGAG 420
DB      424 CGCGTGAAGTCAGAGAACAAAGGTAGTAATCTTGACCTCTTTCGACCCGCTCCGAGCGAG 483
QY      421 GAGGATGACCGGGAAAGTGTCCGTCGCGGAGATCTTCGCGAATTCAGAAATTCACA 480
DB      484 GAGGATGACCGGGAAAGTGTCCGTCGCGGAGATCTTCGCGAATTCAGAAATTCACA 543
QY      481 CCAGGATGCCCGCAATGGGCAAGCGCCCGATTCAACCTCCGCTGTGAGTCTTGAAG 540
DB      544 CCAGGATGCCCGCAATGGGCAAGCGCCCGATTCAACCTCCGCTGTGAGTCTTGAAG 603
QY      541 GCCCGGACTAGTCCCTCCAGTGTATACGGGTGCGCACTGCACTTAAGACCCCT 600
DB      604 GCCCGGACTAGTCCCTCCAGTGTATACGGGTGCGCACTGCACTTAAGACCCCT 663
QY      601 CCTATACCACTTCACAGGAGAAAGAGACAGTTGTTCTGACAGATCCACCGTCTTCT 660
DB      664 CCTATACCACTTCACAGGAGAAAGAGACAGTTGTTCTGACAGATCCACCGTCTTCT 723
QY      661 GCCCTGGCGGAGGCTTGCACAAAGGCTTTGGTAGCTCCGACCGTCGCGCGACAGC 720
DB      724 GCCCTGGCGGAGGCTTGCACAAAGGCTTTGGTAGCTCCGACCGTCGCGCGACAGC 783
QY      721 GGCACGGCAACGCGCCCTCTGACCAATCTCTCGACGAGCGGAGACAGATCTGACGTT 780
DB      784 GGCACGGCAACGCGCCCTCTGACCAATCTCTCGACGAGCGGAGACAGATCTGACGTT 843
QY      781 GAGTCGTATCTCCATGCCCCCTTGAAGGGGAGACCGGGGACCCGATCTCAGCGAC 840
DB      844 GAGTCGTATCTCCATGCCCCCTTGAAGGGGAGACCGGGGACCCGATCTCAGCGAC 903
QY      841 GGGTCTGGTCTACCGGTGAGGAGCGCGGTGAGAGAGTGCTGCTCGATGTC 900
DB      904 GGGTCTGGTCTACCGGTGAGGAGCGCGGTGAGAGAGTGCTGCTCGATGTC 963
QY      901 TACACATGACAGGCGCTCTGATCAAGCATGCTGCTGCGAGAGAAAGCAAGTGCCTATC 960
DB      964 TACACATGACAGGCGCTCTGATCAAGCATGCTGCTGCGAGAGAAAGCAAGTGCCTATC 1023
QY      961 AACGCGTTGAGCAACTCTTGTCTGCTGCTACCAACACATGTTAGCTTACCAATCCG 1020
DB      1024 AACGCGTTGAGCAACTCTTGTCTGCTGCTACCAACATGTTAGCTTACCAATCCG 1083
QY      1021 AGCGCAAGCGACGCG 1035
DB      1084 AGCGCAAGCGACGCG 1098

```

RESULT 3  
AAQ12237  
ID AAQ12237 standard; DNA; 1035 BP.  
AC AAQ12237;  
AC 25-MAR-2003 (revised)  
DT 06-SEP-1991 (first entry)  
XX Clone U3 encoding PT-NANBH virus antigenic portion.  
XX post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.  
XX Non-A.  
XX non-B hepatitis virus.  
XX GB2239245-A.  
XX PN 26-JUN-1991.  
XX PD  
XX

PF 17-DEC-1990; 90GB-00027250.  
 XX 18-DEC-1989; 89GB-00028562.  
 XX 27-FEB-1990; 90GB-00004414.  
 PR 03-MAR-1990; 90GB-00004814.  
 XX  
 PA (WELL ) WELLCOME FOUNDT LTD.  
 PA (HIGH/) HIGHFIELD P E.  
 XX  
 PI Highfield PE, Rodgers BC, Tedder RS, Barbara JAJ;  
 XX  
 DR WPI, 1991-187584/26.  
 XX P-PSDB; AAR12595.  
 PT Post-transferronal non-A non-B hepatitis poly:peptide(s) - and also DNA  
 XX and antibodies used in diagnostic assays and in vaccines.  
 XX  
 PS Claim 10; Page 53-55; 108pp; English.  
 CC This sequence is a non-structural region of the PT-NANBH viral genome  
 CC encoding an antigenic polypeptide. It was isolated from serum of humans  
 CC infected by the virus. Genomic RNA from viruses pelleted from the  
 CC infected serum was used to prepare a cDNA library in lambda stII which  
 CC was screened with antibodies from the original serum and oligonucleotide  
 CC probes. One of the positive plaques to be obtained was J63. DNA was  
 CC extracted from denatured phage stocks, amplified by PCR and sequenced.  
 CC See also AAQ12236 and AAQ12238-Q12242. (Updated on 25-MAR-2003 to correct  
 CC PA field.)  
 CC  
 SQ Sequence 1035 BP, 212 A; 340 C; 296 G; 187 T; 0 U; 0 Other;

Query Match 99.8%; Score 1033.4; DB 2; Length 1035;  
 Best Local Similarity 99.9%; Pred. No. 5.2e-268;  
 Matches 1034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGAAAGTGGATGGGCTGGCTGCAAGGTACGCTCCGGCGTGAACCTCTCTACGG 60  
 DB 1 AAGAAAGTGGATGGGCTGGCTGCAAGGTACGCTCCGGCGTGAACCTCTCTACGG 60  
 QY 61 GAGGAGGTGACATTCAGGTCGGGGCTCAACCAATACCTGGTGGGTCGAGTCCCATGC 120  
 DB 61 GAGGAGGTGACATTCAGGTCGGGGCTCAACCAATACCTGGTGGGTCGAGTCCCATGC 120  
 QY 121 GAGCCCGAACCAGATGTAGCACTGCTCACTTCATGCTACCGACCCCTCCCATACCA 180  
 DB 121 GAGCCCGAACCAGATGTAGCACTGCTCACTTCATGCTACCGACCCCTCCCATACCA 180  
 QY 121 GAGTCCGAACCGGATGTAGCACTGCTCACTTCATGCTACCGACCCCTCCCATACCA 180  
 DB 121 GAGTCCGAACCGGATGTAGCACTGCTCACTTCATGCTACCGACCCCTCCCATACCA 180  
 QY 181 GCAAGACCGGCTTAAGCGCAGGCTGGCCAGGGGGTCTCCCTCTGGCCAGCTCTTCA 240  
 DB 181 GCAAGACCGGCTTAAGCGCAGGCTGGCCAGGGGGTCTCCCTCTGGCCAGCTCTTCA 240  
 QY 241 GCTAGCCAGTTGTCTGGCCCTCTCGAAGGGGAGCATATATACCAAAATATCTTCCA 300  
 DB 241 GCTAGCCAGTTGTCTGGCCCTCTCGAAGGGGAGCATATATATACCAAAATATCTTCCA 300  
 QY 301 GACGCTGACCTCATGAGGCGCAACCTCTGTGGCGCATGAGATGGGCGGGACATTACC 360  
 DB 301 GACGCTGACCTCATGAGGCGCAACCTCTGTGGCGCATGAGATGGGCGGGACATTACC 360  
 QY 361 GCGGTGAGTCAAGAAACAAGTATATCTTGACCTCTTTCAGACCCGCTCCGAGCGAG 420  
 DB 361 GCGGTGAGTCAAGAAACAAGTATATCTTGACCTCTTTCAGACCCGCTCCGAGCGAG 420  
 QY 421 GAGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTCGGGAATCCAAAGAAATTTCCA 480  
 DB 421 GAGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTCGGGAATCCAAAGAAATTTCCA 480  
 QY 481 CCAAGGATGCCCCGATGGGACAGCCCGGATTAACAACCTCCGCTGTGAGTCTGGAAG 540  
 DB 481 CCAAGGATGCCCCGATGGGACAGCCCGGATTAACAACCTCCGCTGTGAGTCTGGAAG 540  
 QY 541 GCCCGGAGTACGTCCCTCAAGTGTACATGGGTCCCACTGACCTTAAGAACCCT 600  
 DB 541 GCCCGGAGTACGTCCCTCAAGTGTACATGGGTCCCACTGACCTTAAGAACCCT 600

DB 541 GCCCGGAGTACGTCCCTCAAGTGTACATGGGTCCCACTGACCTTAAGAACCCT 600  
 QY CCTATACCACTCCACGGAGAAAGAGCAGTTGTTTGAAGAAATCCACCGTGTCTCT 660  
 DB CCTATACCACTCCACGGAGAAAGAGCAGTTGTTTGAAGAAATCCACCGTGTCTCT 660  
 QY 661 GCCCTGGGAGCTTTCACAAAGCTTTTGTAGCTTCGAGACCGTCCGCGGACAGC 720  
 DB 661 GCCCTGGGAGCTTTCACAAAGCTTTTGTAGCTTCGAGACCGTCCGCGGACAGC 720  
 QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCCAGCAGCGCGAGAGATCTGAGCTT 780  
 DB 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCCAGCAGCGCGAGAGATCTGAGCTT 780  
 QY 781 GAGTGTATTCCTCATGCCCCCTTGAAGGGAGAGCCGGGAGACCTCCGATCTGAGCAG 840  
 DB 781 GAGTGTATTCCTCATGCCCCCTTGAAGGGAGAGCCGGGAGACCTCCGATCTGAGCAG 840  
 QY 841 GGGTCTTGATTCACCGTAGTAGAGGCGGTGAGACGTCGTCTGCTCGATGTC 900  
 DB 841 GGGTCTTGATTCACCGTAGTAGAGGCGGTGAGACGTCGTCTGCTCGATGTC 900  
 QY 901 TACATGAGACAGGCGCTCTGATCAAGCCATGCGCTGGGAGAAAGCAAGCTGCCATC 960  
 DB 901 TACATGAGACAGGCGCTCTGATCAAGCCATGCGCTGGGAGAAAGCAAGCTGCCATC 960  
 QY 961 AACGGTGAACAACCTTTGTGCTGTCACCAACATGCTACGCTACCAATCCCGC 1020  
 DB 961 AACGGTGAACAACCTTTGTGCTGTCACCAACATGCTACGCTACCAATCCCGC 1020  
 QY 1021 AGCGCAAGCCAGCG 1035  
 DB 1021 AGCGCAAGCCAGCG 1035  
 RESULT 4  
 AAQ46193  
 ID AAQ46193 standard; DNA; 1788 BP.  
 XX  
 AC AAQ46193;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 24-FEB-1994 (first entry)  
 XX  
 DE PT-NANBH virus NS5 protein.  
 XX  
 KW Parenterally transmitted non A non B hepatitis; PT-NANBH;  
 KW hepatitis C virus; HCV; detection; diagnosis; antigen; vaccine; ss.  
 OS Hepatitis virus.  
 XX  
 FN W09317110-A2.  
 XX  
 PD 02-SEP-1993.  
 XX  
 PF 19-FEB-1993; 93WO-GB000345.  
 XX  
 PR 21-FEB-1992; 92GB-00003803.  
 XX  
 PA (WELL ) WELLCOME FOUNDT LTD.  
 XX  
 PI Parker D, Rodgers BC;  
 XX  
 DR WPI, 1993-288415/36.  
 DR P-PSDB; AAR1433.  
 XX  
 PT New recombinant polypeptide for diagnosing hepatitis C - contains three  
 PT distinct antigens from different viral regions, also useful in protective  
 XX  
 PS Claim 7; Page 33-36; 99pp; English.  
 XX

CC If at least three different PT-NANBH antigens are used to screen for PT-NANBH, the screening is much more sensitive as compared to the use of CC only two PT-NANBH antigens. Pref. antigens are described in AAQ46192-94. CC Two new antigenic regions of the PT-NANBH genome are given in AAQ46198-99. AAQ46202 describes an improved PT-NANBH recombinant polypeptide. CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 1788 BP; 401 A; 544 C; 498 G; 345 T; 0 U; 0 Other;

Query Match 99.8%; Score 1033.4; DB 2; Length 1788;

Best Local Similarity 99.9%; Pred. No. 6.2e-268;

Matches 1034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ACAGAAAGTGGATGGGCTGCGGCTGCAAGGTACGCTTCGGCGTGCAGAACTCTCTTACCG 60
DB 1 ACAGAAAGTGGATGGGCTGCGGCTGCAAGGTACGCTTCGGCGTGCAGAACTCTCTTACCG 60
QY 61 GAGGAGGTGCAATTCAGGCTCGGCTCAACCAATACCTGGTGGGTGCGACGCTCCCATGC 120
DB 61 GAGGAGGTGCAATTCAGGCTCGGCTCAACCAATACCTGGTGGGTGCGACGCTCCCATGC 120
QY 61 GAGGAGGTGCAATTCAGGCTCGGCTCAACCAATACCTGGTGGGTGCGACGCTCCCATGC 120
DB 61 GAGGAGGTGCAATTCAGGCTCGGCTCAACCAATACCTGGTGGGTGCGACGCTCCCATGC 120
QY 121 GAGCCCGAACCAGGATAGCAGTGTCTCACTTCCATGCTCAACGACCCCTCCCATCACA 180
DB 121 GAGCCCGAACCAGGATAGCAGTGTCTCACTTCCATGCTCAACGACCCCTCCCATCACA 180
QY 121 GAGCCCGAACCAGGATAGCAGTGTCTCACTTCCATGCTCAACGACCCCTCCCATCACA 180
DB 121 GAGCCCGAACCAGGATAGCAGTGTCTCACTTCCATGCTCAACGACCCCTCCCATCACA 180
QY 181 GCAAGACCGCTTAAGCGCAGGCTGCGCAGAGGGGTCTCCCTCTTGGCGACGCTTCTCA 240
DB 181 GCAAGACCGCTTAAGCGCAGGCTGCGCAGAGGGGTCTCCCTCTTGGCGACGCTTCTCA 240
QY 241 GCTAGCAATGTGCTGCGGCTCTCTGGAAGGCGACATTAATCCCAAAATGATCTTCCCA 300
DB 241 GCTAGCAATGTGCTGCGGCTCTCTGGAAGGCGACATTAATCCCAAAATGATCTTCCCA 300
QY 301 GACGCTGACCTGATCGAGGCGCAACTCTCTGTGCGCGCATGAGATGGGCGGAGCATTAAC 360
DB 301 GACGCTGACCTGATCGAGGCGCAACTCTCTGTGCGCGCATGAGATGGGCGGAGCATTAAC 360
QY 361 GCGGTGAGTCAAGAGAACAGATGTAATCTCTGACTCTTTTGAACCCGCTCCGAGCGAG 420
DB 361 GCGGTGAGTCAAGAGAACAGATGTAATCTCTGACTCTTTTGAACCCGCTCCGAGCGAG 420
QY 421 GAGGATGAGGAGGAGTGTCCGTCGCGGAGGAGATCTCTCGGAAATCCAAAGAAATCCCA 480
DB 421 GAGGATGAGGAGGAGTGTCCGTCGCGGAGGAGATCTCTCGGAAATCCAAAGAAATCCCA 480
QY 481 CGAGCATGCCGATGAGGCGACGCCGAGTATCAACCTCTCCGCTGCTGAGTCTTGAAG 540
DB 481 CGAGCATGCCGATGAGGCGACGCCGAGTATCAACCTCTCCGCTGCTGAGTCTTGAAG 540
QY 541 GCGCCGAGTACGCTCCCTCAAGTGTATCATGGGTGCCACTGCGACCTTACTAAGACCCCT 600
DB 541 GCGCCGAGTACGCTCCCTCAAGTGTATCATGGGTGCCACTGCGACCTTACTAAGACCCCT 600
QY 601 CCTATATCACTTCCAGGAGAAAGAGAGATGTTCTTGAAGAAATCCACCGGTCTTCT 660
DB 601 CCTATATCACTTCCAGGAGAAAGAGAGATGTTCTTGAAGAAATCCACCGGTCTTCT 660
QY 661 GCGCTGCGGAGCTTGCACAAAGCTTTGTGATGCTCCGAGACCGTCCGCGCTGACAGC 720
DB 661 GCGCTGCGGAGCTTGCACAAAGCTTTGTGATGCTCCGAGACCGTCCGCGCTGACAGC 720
QY 721 GGCACGCGAACCGCCCTCTCTGACCAATCTCTCGACGACGCGCGAGACAGATCTGACGT 780
DB 721 GGCACGCGAACCGCCCTCTCTGACCAATCTCTCGACGACGCGCGAGACAGATCTGACGT 780
QY 781 GAGTGTATTTCTTCATATGCCCCCTTGAAGGGGAGACCCGAGATCTCAACGAC 840
DB 781 GAGTGTATTTCTTCATATGCCCCCTTGAAGGGGAGACCCGAGATCTCAACGAC 840
QY 841 GGGTCTTGGTCTACCGTGAAGTGAAGGCGGTGAAGACCTGCTGCTGCTGATGTCC 900
DB 841 GGGTCTTGGTCTACCGTGAAGTGAAGGCGGTGAAGACCTGCTGCTGCTGATGTCC 900

```

```

QY 901 TACACATGAGCAGGCGCTGTGATCAAGCCATGCGCTGCGGAGGAAAGCAAGCTGCCATC 960
DB 901 TACACATGAGCAGGCGCTGTGATCAAGCCATGCGCTGCGGAGGAAAGCAAGCTGCCATC 960
QY 961 AACGCGTTGAGCAACTTTTGTGCGTCAACCAACATGATGCTACGCTACCATCCGC 1020
DB 961 AACGCGTTGAGCAACTTTTGTGCGTCAACCAACATGATGCTACGCTACCATCCGC 1020
QY 1021 ACGGCAAGCCAGCGG 1035
DB 1021 ACGGCAAGCCAGCGG 1035

```

# RESULT 5

AAQ12241  
ID AAQ12241 standard; cDNA; 3750 BP.

XX AAQ12241;

AC 25-MAR-2003 (revised)

DT 17-SEP-1991 (first entry)

XX Encodes portion of PT-NANBH viral non-structural protein.

KW post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.

OS Non-A.

XX non-B hepatitis virus.

PN GB2239245-A.

XX 26-JUN-1991.

XX 17-DEC-1990; 90GB-00027250.

XX 18-DEC-1989; 89GB-00028562.

PR 27-FEB-1990; 90GB-00004414.

PR 03-MAR-1990; 90GB-00004814.

XX (WELL) WELLCOME FOUND LTD.

PA (HIGH) HIGHFIELD P E.

PI Highfield PE, Rodgers BC, Tedder RS, Barbara JAJ;

DR WPI; 1991-187584/26.

XX P-PSDB; AAR12599.

XX Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA

PT and antibodies used in diagnostic assays and in vaccines.

XX Claim 10; Page 88-97; 108pp; English.

CC This sequence probably encodes viral non-structural proteins of the PT-NANBH viral genome which are antigenic. It was isolated from serum of humans infected by the virus. See also AAQ12236-40 and AAQ12242. (Updated on 25-MAR-2003 to correct PA field.)

CC Sequence 3750 BP; 794 A; 1140 C; 1072 G; 744 T; 0 U; 0 Other;

Query Match 99.1%; Score 1025.4; DB 2; Length 3750;

Best Local Similarity 99.4%; Pred. No. 1.1e-265;

Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 ACAGAAAGTGGATGGGCTGCGGCTGCAAGGTACGCTTCGGCGTGCAGAACTCTCTTACCG 60
DB 1696 ACAGAAAGTGGATGGGCTGCGGCTGCAAGGTACGCTTCGGCGTGCAGAACTCTCTTACCG 1755
QY 61 GAGGAGTCAATTCAGGCTCGGCTCAACCAATACCTGGTGGGTGCGACGCTCCCATGC 120
DB 1756 GAGGAGTCAATTCAGGCTCGGCTCAACCAATACCTGGTGGGTGCGACGCTCCCATGC 1815
QY 121 GAGCCCGAACCAGGATAGCAGTGTCTCACTTCCATGCTCAACGACCCCTCCCATCACA 180

```

```

Db      1816 GAGCCGAAACCGGATGTAGACAGTGTCACTTCCATGTCTACCGACCCCTCCCACTACACA      1875
Qy      181  GAGAGACCGGCTAAAGCGAGGCTGGCCAGAGGGGTCTCCCCCTCTCTTGCCAGCTCTTCA      240
Db      1876 GAGAGACCGGCTAAAGCGAGGCTGGCCAGAGGGGTCTCCCCCTCTCTTGCCAGCTCTTCA      1935
Qy      241  GCTAGCCAGATTGCTTGAGCCCTTCTCTGAAAGGCGACATACATTACCMAAATGACTTCCCA      300
Db      1936 GCTAGCCAGATTGCTTGAGCCCTTCTCTGAAAGGCGACATACATTACCMAAATGACTTCCCA      1995
Qy      301  GACGCTGACCTCATGAGGCGCAACCTCTGTGGCGGATGAGATGGGCGGAGCAATTAC      360
Db      1996 GACGCTGACCTCATGAGGCGCAACCTCTGTGGCGGATGAGATGGGCGGAGCAATTAC      2055
Qy      361  CGCGTGAGTCAAGAAACAAGGTAGTAATCTGTGACTCTTTTGACCCCTCCGACCGGAG      420
Db      2056 CGCGTGAGTCAAGAAACAAGGTAGTAATCTGTGACTCTTTTGACCCCTCCGACCGGAG      2115
Qy      421  GAGATGAGCGGGAAGTGTCCGTCCCGGCGAGATCTCTCGGAAATCCAAAAATTCCCA      480
Db      2116 GAGATGAGCGGGAAGTGTCCGTCCCGGCGAGATCTCTCGGAAATCCAAAAATTCCCA      2175
Qy      481  CCAAGCATGCCCGCATGGGCAAGCCCGGATTAACAACCTTCGCTGTGAGTCTTGGAAG      540
Db      2176 CCAAGCATGCCCGCATGGGCAAGCCCGGATTAACAACCTTCGCTGTGAGTCTTGGAAG      2235
Qy      541  GCGCCGGATAGTGTCCCTCCATGTGTACATGGGTGCCCATGCGCACTTAATAAGACCCCT      600
Db      2236 GCGCCGGATAGTGTCCCTCCATGTGTACATGGGTGCCCATGCGCACTTAATAAGACCCCT      2295
Qy      601  CCTATACACACTTCAACGAGAAAGAGACAGTGTCTGTACAGAAATCCACCGTCTTCT      660
Db      2296 CCTATACACACTTCAACGAGAAAGAGACAGTGTCTGTACAGAAATCCACCGTCTTCT      2355
Qy      661  GCGCTGGCGAGCTTGCCAAAGGCTTGTGTAGTCTCGAACCGTGGCGCTGCAACAGC      720
Db      2356 GCGCTGGCGAGCTTGCCAAAGGCTTGTGTAGTCTCGAACCGTGGCGCTGCAACAGC      2415
Qy      721  GGCACGGGCAACCGCCCTCTCAACCAATCTCCGACGAGGCGGAGAGAGATCTGACGT      780
Db      2416 GGCACGGGCAACCGCCCTCTCAACCAATCTCCGACGAGGCGGAGAGAGATCTGACGT      2475
Qy      781  GAGTCGATTCCTCCATGCCCCCTTGAAGGGGAGCGCGGAGACCCCGATCTCAGCGAC      840
Db      2476 GAGTCGATTCCTCCATGCCCCCTTGAAGGGGAGCGCGGAGACCCCGATCTCAGCGAC      2535
Qy      841  GGGTCTTGCTTACCGGTGAGTGAAGAGCGCGGTGAGAGACGTGCTGTGCTGATGTCC      900
Db      2536 GGGTCTTGCTTACCGGTGAGTGAAGAGCGCGGTGAGAGACGTGCTGTGCTGATGTCC      2595
Qy      901  TACACATGAGCAAGGGCTGTGATCAAGCATGCGCTGCGAGAAAGCAAGTGGCCATC      960
Db      2596 TACACATGAGCAAGGGCTGTGATCAAGCATGCGCTGCGAGAAAGCAAGTGGCCATC      2655
Qy      961  AACGGTTGAGCAACTTGTGTGCTGACCAACAATGCTTACGCTTACCAATCCGC      1020
Db      2656 AACGGTTGAGCAACTTGTGTGCTGACCAACAATGCTTACGCTTACCAATCCGC      2715
Qy      1021 AAGCGCAAGCCAGCGG      1035
Db      2716 AAGCGCAAGCCAGCGG      2730

```

```

RESULT 6
AA046195
ID      AA046195 standard; cDNA to mRNA; 7065 BP.
AC      AA046195;
XX
XX      27-AUG-2003 (revised)
DT      25-MAR-2003 (revised)
DT      24-FEB-1994 (first entry)

```

```

XX      PT-NANBH virus non-structural proteins.
DE
XX      Parenterally transmitted non A non B hepatitis; PT-NANBH; NS4;
KM      hepatitis C virus; HCV; detection; diagnosis; antigen; vaccine;
KM      amplification; primer; polymerase chain reaction; PCR; ss.
XX      Hepatitis virus.
XX      OS
PN      WO9317110-A2.
XX
XX      02-SEP-1993.
PD
XX      19-FEB-1993; 93WO-GB000345.
PF
XX      21-FEB-1992; 92GB-00003803.
PR
XX      (WELL ) WELLCOME FOUND LTD.
PA
XX      Parker D, Rodgers BC;
XX
XX      WPI: 1993-288415/36.
DR      P-FSDB; AAR41435.
DR
XX      New recombinant polypeptide for diagnosing hepatitis C - contains three
PT      distinct antigens from different viral regions, also useful in protective
PT      vaccines.
XX
XX      Example 1; Page 43-53; 99pp; English.
PS
XX
XX      The NS4 region from the 3' region of the PT-NANBH genome (AA046195) is
CC      amplified by PCR using primers D224 and D226 (AA046196-97) and the
CC      fragment (AA046198) is cloned into a vector and expressed in infected
CC      insect cells. The recombinant virus (BHC-13) was able to express the NS4
CC      specific recombinant protein at low levels in the infected insect cells.
CC      If at least three different PT-NANBH antigens are used to screen for PT-
CC      NANBH, the screening is much more sensitive as compared to the use of
CC      only two PT-NANBH antigens. Pref. antigens are described in AA046192-94.
CC      Two new antigenic regions of the PT-NANBH genome are given in AA046198-
CC      99. AA046202 describes an improved PT-NANBH recombinant polypeptide.
CC      (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to
CC      correct OS field.)
XX
XX      Sequence 7065 BP; 1468 A; 2123 C; 1994 G; 1480 T; 0 U; 0 Other;
SQ

```

```

Query Match      99.1%; Score 1025.4; DB 2; Length 7065;
Best Local Similarity 99.4%; Pred. No. 1.4e-265;
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy      1  ACAGAGTGAATGGGGTGGCTGCTGACAGTACGCTCCGGCGTGCMAAATCTCTCTACGG      60
Db      3919 ACAGAGTGAATGGGGTGGCTGCTGACAGTACGCTCCGGCGTGCMAAATCTCTCTACGG      3978
Qy      61  GAGAGGTCACTTCAGAGTGGGCTCAACCAATACCTGTGTGGTGGTGGAGCTCCCATGC      120
Db      3979 GAGAGGTCACTTCAGAGTGGGCTCAACCAATACCTGTGTGGTGGTGGAGCTCCCATGC      4038
Qy      121  GAGCCCGAAACGGATGTAGAGTGTGCTCACTTCCATGTGTCAACCGACCCCTCCCAATACCA      180
Db      4039 GAGCCCGAAACGGATGTAGAGTGTGCTCACTTCCATGTGTCAACCGACCCCTCCCAATACCA      4098
Qy      181  GCAAGACGCGCTAAGCGCAGGCTGGCGAGGGGTCTCCCCCTCTCTTGCCAGCTCTTCA      240
Db      4099 GCAAGACGCGCTAAGCGCAGGCTGGCGAGGGGTCTCCCCCTCTCTTGCCAGCTCTTCA      4158
Qy      241  GCTAGCCAGTTGTCTGCGCCCTTCTCTGAAAGCGACATACATTACCAAAATGACTTCCCA      300
Db      4159 GCTAGCCAGTTGTCTGCGCCCTTCTCTGAAAGCGACATACATTACCAAAATGACTTCCCA      4218
Qy      301  GACGCTGACCTCATGAGGCGCAACCTCTGTGTGGCGGATGAGATGGGCGGAGCAATTACC      360
Db      4219 GACGCTGACCTCATGAGGCGCAACCTCTGTGTGGCGGATGAGATGGGCGGAGCAATTACC      4278

```



QY 361 CCGGTGAGTACAGAAAGAGTAGTAACTCTGAGCTCTTTGACCCGCTCGAGCGAG 420  
DB 4279 CCGGTGAGTACAGAAAGAGTAGTAACTCTGAGCTCTTTGACCCGCTCGAGCGAG 4338  
QY 421 GAGGATGACGGGAAAGTGTCCGTCCCGCGAGAGATCTTCGGAAATTCAGAAATTTCCCA 480  
DB 4339 GAGGATGACGGGAAAGTGTCCGTCCCGCGAGAGATCTTCGGAAATTCAGAAATTTCCCA 4398  
QY 481 CCAAGCATCCCGCATGGGCAAGCCCGGATTAACACCTTCGCTGAGTCTTGGAG 540  
DB 4399 CCAAGCATCCCGCATGGGCAAGCCCGGATTAACACCTTCGCTGAGTCTTGGAG 4458  
QY 541 GCCCGGAGTACGTCCTCCAGTGTATACATGGTGGCCCACTGCGACCTACTAAGACCCCT 600  
DB 4459 GCCCGGAGTACGTCCTCCAGTGTATACATGGTGGCCCACTGCGACCTACTAAGACCCCT 4518  
QY 601 CCTATACACCTTCACGAGAAAGAGACAGTTGTTCTGACAGAAATTCACCGTGTCTCT 660  
DB 4519 CCTATACACCTTCACGAGAAAGAGACAGTTGTTCTGACAGAAATTCACCGTGTCTCT 4578  
QY 661 GCGCTGCGGAGTGTCCCAAAAGCTTTTGTAGTCTCCGAGACCTGCGCTCGACAGC 720  
DB 4579 GCGCTGCGGAGTGTCCCAAAAGCTTTTGTAGTCTCCGAGACCTGCGCTCGACAGC 4638  
QY 721 GGGACGGCAACCGCCCTCTGACCAATCTCCGAGACCGCGAGAGATCTGACGTT 780  
DB 4639 GGGACGGCAACCGCCCTCTGACCAATCTCCGAGACCGCGAGAGATCTGACGTT 4698  
QY 781 GAGTGTATTCCTTCATGCCCCCTTGAAGGGGAGCCGGGAGACCCGATCTCAAGGAC 840  
DB 4699 GAGTGTATTCCTTCATGCCCCCTTGAAGGGGAGCCGGGAGACCCGATCTCAAGGAC 4758  
QY 841 GGGTCTTGTGTCTACCGTGAAGAGAGCGGTGAGAGACGTCGTGCTGCTGATGTCC 900  
DB 4759 GGGTCTTGTGTCTACCGTGAAGAGAGCGGTGAGAGACGTCGTGCTGCTGATGTCC 4818  
QY 901 TACACATGAGCAGGCGCTGTATCAGCCCAATGCGCTGCGGAGAAAGCAAGTCCCATC 960  
DB 4819 TACACATGAGCAGGCGCTGTATCAGCCCAATGCGCTGCGGAGAAAGCAAGTCCCATC 4878  
QY 961 AAGCGCTTGAACACTTTTGTGCTGCTACACCAAGATGTTCTACGCTACCATCCCGC 1020  
DB 4879 AAGCGCTTGAACACTTTTGTGCTGCTACACCAAGATGTTCTACGCTACCATCCCGC 4938  
QY 1021 AAGCGCAAGCCAGCGG 1035  
DB 4939 AAGCGCAAGCCAGCGG 4953

RESULT 7  
ADD93733 standard; DNA; 7989 BP.  
ID ADD93733;  
AC ADD93733;  
XX 29-JAN-2004 (first entry)  
DT 29-JAN-2004 (first entry)  
XX Hepatitis C virus strain J4 (BB7/JANSSB) replicon.  
DE Hepatitis C virus; virucide; ss.  
XX HCV; vaccine; virucide; ss.  
OS Hepatitis C virus.  
XX WO2003085084-A2.  
PN 16-OCT-2003.  
XX 03-APR-2003; 2003WO-US010177.  
PF 03-APR-2003; 2002US-036985P.  
PR 03-APR-2002; 2002US-036985P.  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
XX

PI Gates A, Gu B, Sarsky RT;  
XX WPI, 2003-804301/75.  
XX New hepatitis C virus (HCV) sub-genomic replicon, useful for facilitating  
PT screening or testing of anti-HCV drugs, comprises a nucleic acid  
PT construct encoding chimeric HCV non-structural proteins, and an NS5B  
PT polymerase gene.  
XX disclosure; Page 80-85; 159pp; English.  
XX  
XX The present sequence comprises a replicating hepatitis C virus (HCV) J4  
CC (BB7/JANSSB) replicon. The invention provides sub-genomic replicons of  
CC HCV comprising a nucleic acid construct encoding chimeric HCV  
CC nonstructural protein and an NS5B polymerase gene. A preferred replicon  
CC comprises an NS3 nucleotide sequence ADD93721 that encodes the first 75  
CC contiguous N-terminal amino acids of the NS3 of genotype 1b, of a BB7  
CC strain. A chimeric replicon may comprise an NS3 sequence from any of the  
CC 6 major HCV genotypes and subtypes but has its first 225 nucleotides of  
CC the coding sequence replaced by the BB7 strain NS3 sequence, especially  
CC where the replicon is from HCV genotype 1a (H77 strain) or genotype 1b  
CC (J4 strain). Stable cell lines expressing and replicating functional  
CC replicons containing sequences from HCV genotype 1a (strain H77) or  
CC genotype 1b (strain J4) within the prototype 1b replicon backbone from  
CC HCV strain BB7 are provided. These can be used to screen for compounds  
CC that modulate viral replication. The sub-genomic HCV replicon systems of  
CC the invention may provide the foundation for generating HCV replicons of  
CC all 6 major genotypes and subtypes to facilitate screening, testing and  
CC evaluating anti-infective agents for HCV disease(s).  
XX  
SQ Sequence 7989 BP; 1656 A; 2373 C; 2230 G; 1730 T; 0 U; 0 Other;  
Query Match 85.9%; Score 889.4; DB 10; Length 7989;  
Best Local Similarity 91.2%; Pred. No. 6.1e-229;  
Matches 944; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
QY 1 ACAGAGTGTATGAGGAGTGGCGGTGACAGATAGCTCCGCGGACCAACTCTCTACGG 60  
DB 5092 ACAGAGTGTATGAGGAGTGGCGGTGACAGATAGCTCCGCGGACCAACTCTCTACGG 5151  
QY 61 GAGGAGTGTATGAGGAGTGGCGGTGACAGATAGCTCCGCGGACCAACTCTCTACGG 120  
DB 5152 GAGGAGTGTATGAGGAGTGGCGGTGACAGATAGCTCCGCGGACCAACTCTCTACGG 5211  
QY 121 GAGCCCGAACCAGATGTACAGTGTCTCACTTTCATCTCAACGACCCCTCCCATCA 180  
DB 5212 GAGCCCGAACCAGATGTACAGTGTCTCACTTTCATCTCAACGACCCCTCCCATCA 5271  
QY 181 GCAGAGAGCGGTAAAGCGAGGCTGGCGAGGAGGTCTCCCGCTCTTGGCAAGCTTTCA 240  
DB 5272 GCAGAGAGCGGTAAAGCGAGGCTGGCGAGGAGGTCTCCCGCTCTTGGCAAGCTTTCA 5331  
QY 241 GCTAGCAGTGTGTGCGCCCTTCTCTGAGAGGAGCAATACATTACCAAAATGACTTCCA 300  
DB 5332 GCTAGCAGTGTGTGCGCCCTTCTCTGAGAGGAGCAATACATTACCAAAATGACTTCCA 5391  
QY 301 GACGCTGACCTCATGAGGCGCAACTCTCTGTGGCGGAGATGAGTGGCGGAGATTAC 360  
DB 5392 GACGCTGACCTCATGAGGCGCAACTCTCTGTGGCGGAGATGAGTGGCGGAGATTAC 5451  
QY 361 CCGGTGAGTACAGAAAGAGTAGTAACTCTGAGCTCTTTGACCCGCTCGAGCGAG 420  
DB 5452 CCGGTGAGTACAGAAAGAGTAGTAACTCTGAGCTCTTTGACCCGCTCGAGCGAG 5511  
QY 421 GAGGATGACGGGAAAGTGTCCGTCCCGCGAGAGATCTTCGGAAATTCAGAAATTTCCCA 480  
DB 5512 GAGGATGACGGGAAAGTGTCCGTCCCGCGAGAGATCTTCGGAAATTCAGAAATTTCCCA 5571  
QY 481 CCAAGCATCCCGCATGGGCAAGCCCGGATTAACACCTTCGCTGAGTCTTGGAG 540  
DB 5572 CCAAGCATCCCGCATGGGCAAGCCCGGATTAACACCTTCGCTGAGTCTTGGAG 5631  
QY 541 GCCCGGAGTACGTCCTCCAGTGTATACATGGTGGCCCACTGCGACCTACTAAGACCCCT 600





QY 601 CCTATACCACTTCCACGAGAAAAGAGACAGTTGTTCTGACAGAAATCCACCGTGTCTTCT 660  
DB 5692 CCGATPACCACTTCCACGAGAAAAGAGACAGTTGTTCTGACAGAAATCCACCGTGTCTTCT 5751  
QY 661 GCGCTGGCGGAGCTTGGCCAAAGGCTTTTGTAGTCCGGACCGTGGCCGCTCCAGAC 720  
DB 5752 GCGTGGCGGAGCTTGGCCAAAGGCTTTTGTAGTCCGGACCGTGGCCGCTCCAGAC 5811  
QY 721 GGCACGGCAACGCGCCCTCTGACCAATCTCCGACGACGCGGAGCAGATCTGACGTT 780  
DB 5812 GGCACGGCAACGCGCCCTCTGACCAAGCCCTCCGACGACGCGGAGCAGATCTGACGTT 5871  
QY 781 GAGTCGTATTCCTCAATGCCCCCTTGAAGGAGGAGCCGAGGACCCCGATTCAGCGAC 840  
DB 5872 GAGTCGTATTCCTCAATGCCCCCTTGAAGGAGGAGCCGAGGATCCCGATTCAGCGAC 5931  
QY 841 GGGTCTGGTTCACCGTAGTAGAGAGCGCGGTGAGAGAGCTGTCTGCTCTGATGTC 900  
DB 5932 GGGTCTGGTTCACCGTAGTAGAGAGCGCGGTGAGAGAGCTGTCTGCTGATGTC 5991  
QY 901 TACACATGACAGAGCGCTGTGATCAAGCCATGCGCTGAGGAGAAAGCAAGCTGCCATC 960  
DB 5992 TACACATGACAGAGCGCGCTGTGATCAAGCCATGCGCTGAGGAGAAAGCAAGCTGCCATC 6051  
QY 961 AACGCGTGAACAATCTTTGCTGCGTCAACCAACATGCTTACGCTACCAATCCGCG 1020  
DB 6052 AATGCACTGAGCAACTTTGCTGCGTCAACCAACTTGTCTATGCTACCAATCTGCG 6111  
QY 1021 AGCGCAAGCGACGCG 1035  
DB 6112 AGCGCAAGCGTGGG 6126

RESULT 10  
AAA98968  
ID AAA98968 strand; DNA; 7989 BP.  
XX  
AC AAA98968;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Hepatitis C virus DNA fragment SEQ ID NO: 4.  
XX  
KW Cell culture; therapy; infection; vaccine; diagnosis; gene therapy; ds.  
XX  
OS Hepatitis C virus.  
XX  
PN DE19915178-A1.  
XX  
PD 05-OCT-2000.  
XX  
PF 03-APR-1999; 99DE-01015178.  
XX  
PR 03-APR-1999; 99DE-01015178.  
XX  
PA (UYMA-) UNIV MAINZ GUTENBERG JOHANNES.  
XX  
PI Bartenschlager R;  
XX  
DR WPI, 2000-629140/61.  
XX  
PT Cell culture system for hepatitis C virus, useful e.g. in screening for  
PT therapeutic agents, comprises human hepatoma cells containing a viral RNA  
PT construct that includes a selectable gene.  
XX  
PS Claim 8; Page 37-43; 58pp; German.  
XX  
CC This invention describes a novel Hepatitis C virus (HCV) cell culture  
CC system comprising human hepatoma cells that contain an integrated HCV-RNA  
CC construct (i). (ii) contains the HCV-specific RNA segments 5'-NTR (non-  
CC translated region), NS (non-structural)1,3, NS4A, NS4B, NS5A, NS5B and 3'-  
CC NTR, and a selectable (marker) gene (iii). The cell cultures, and/or (i),  
CC are used to prepare, evaluate and/or test therapeutic and/or diagnostic

CC agents for HCV infections, and to prepare vaccines against HCV infection  
CC (particularly preparation of attenuated HCV). The can also be used for  
CC preparation of a liver-specific delivery system for gene therapy, and to  
CC identify cells permissive for HCV replication. Virus RNA replicates  
CC autonomously and with high efficiency in this cellular system, so that  
CC variations in replication rates can be measured (for screening antiviral  
CC agents) quantitatively or qualitatively, using standard laboratory  
CC equipment. Efficient replication of HCV RNA is only achieved when the  
CC specified RNA segments are present and when the transfected cells are  
CC maintained under permanent selection pressure

XX Sequence 7989 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other;

Query Match 85.5%; Score 884.6; DB 3; Length 7989;

Best Local Similarity 90.9%; Pred. No. 1.2e-227;

Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1 ACAGAGTGGATGGGAGTGGGCTGACAGTACGCTCCGGGAGTGAACCTCTCTACGG 60  
DB 5092 ACAGAGTGGATGGGAGTGGGCTGACAGTACGCTCCAGGAGTGAACCTCTCTACGG 5151  
QY 61 GAGGAGTGCATATTCAGATCGGAGCTCAACCAATACCTGTTGGGTGACGCTCCATGC 120  
DB 5152 GAGGAGTGCATATTCAGATCGGAGCTCAACCAATACCTGTTGGGTGACGCTCCATGC 5211  
QY 121 GAGCCCGAAGCGATGTAAGAGTGTCTCACTTCACTGCTCAACCGAACCCTCCACATCA 180  
DB 5212 GAGCCCGAAGCGATGTAAGAGTGTCTCACTTCACTGCTCAACCGAACCCTCCACATCA 5271  
QY 181 GAGAGAGCGCTAAGCGAGGCTGGCCAGGAGGAGTCTCCCGCTCTTGGCCAGCTTTCA 240  
DB 5272 GCGGAGAGCGCTAAGCGAGGCTGGCCAGGAGGAGTCTCCCGCTCTTGGCCAGCTTTCA 5331  
QY 241 GCTAGCCAGTGTGTGCGGCTTCTCCGAAAGCGACATACATTAACCAATGACTTCCA 300  
DB 5332 GCTAGCCAGTGTGTGCGGCTTCTCCGAAAGCGACATACATTAACCAATGACTTCCA 5391  
QY 301 GACGCTGACCTCATGAGGCGCAACCTCTGTGAGCGGACATGAGTGGGAGCAATTAAC 360  
DB 5392 GACGCTGACCTCATGAGGCGCAACCTCTGTGAGCGGACATGAGTGGGAGCAATTAAC 5451  
QY 361 GCGGTGAGTCAAGAGCAAGAGTGTATCTCTGACTTTTTCAGCCGCTCCAGCGAG 420  
DB 5452 GCGGTGAGTCAAGAGCAAGAGTGTATCTCTGACTTTTTCAGCCGCTCCAGCGAG 5511  
QY 421 GAGGATGAGCGGAGAGTGTCCGTCGCGGAGATCTCGGAGAAATCCAGAAATTTCCA 480  
DB 5512 GAGGATGAGCGGAGAGTGTCCGTCGCGGAGATCTCGGAGAAATTTCCA 5571  
QY 481 CCAGCGATGCCCGATGAGGACAGCCCGGATTAACAACCTCCGCTGTGAGTCTCGAAG 540  
DB 5572 CGAGCGATGCCCGATGAGGACAGCCCGGATTAACAACCTCCGCTGTGAGTCTCGAAG 5631  
QY 541 GCCCGGACTACGTCCTCTCCAGTGTGTACATGGGTGCCACTGCCACTTAAGAACCCT 600  
DB 5632 GACCCGGACTACGTCCTCTCCAGTGTGTACATGGGTGCCACTGCCACTTAAGAACCCT 5691  
QY 601 CCTATACCACTTCCACGAGAAAAGAGCAGTTGTTCTGACAGAAATCCACCGTGTCTTCT 660  
DB 5692 CCGATPACCACTTCCACGAGAAAAGAGCAGTTGTTCTGACAGAAATCCACCGTGTCTTCT 5751  
QY 661 GCGCTGGCGGAGCTTGGCCAAAGGCTTTTGTAGTCCGGACCGTGGCCGCTCCAGAC 720  
DB 5752 GCGCTGGCGGAGCTTGGCCAAAGGCTTTTGTAGTCCGGACCGTGGCCGCTCCAGAC 5811  
QY 721 GGCACGGCAACGCGCCCTCTGACCAATCTCCGACGACGCGGAGCAGATCTGACGTT 780  
DB 5812 GGCACGGCAACGCGCCCTCTGACCAAGCCCTCCGACGACGCGGAGCAGATCTGACGTT 5871  
QY 781 GAGTCGTATTCCTCAATGCCCCCTTGAAGGAGGAGCCGAGGATCCCGATTCAGCGAC 840  
DB 5872 GAGTCGTATTCCTCAATGCCCCCTTGAAGGAGGAGCCGAGGATCCCGATTCAGCGAC 5931

```

QY 841 GGGCTTGGCTCTACCGGATGAGAGAGCGCGGATGAGAGCTGCTGCTGCTGATGTC 900
DB 5932 GGGCTTGGCTCTACCGGATGAGAGAGCGCGGATGAGAGCTGCTGCTGCTGATGTC 9591
QY 901 TACACATGAGACAGCGCGCTGATGATGATGATGATGATGATGATGATGATGATG 960
DB 5992 TACACATGAGACAGCGCGCTGATGATGATGATGATGATGATGATGATGATGATG 6051
QY 961 AACGCTTGAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 6052 AATGCACTGAGCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6111
QY 1021 AGCGCAAGCCAGCGCG 1035
DB 6112 AGCGCAAGCCAGCGCG 6126

RESULT 11
AAD25322
ID AAD25322 standard; cDNA; 7989 BP.
XX
AC AAD25322;
XX
DT 12-MAR-2002 (first entry)
XX
DE Hepatitis C virus (HCV) replBartman/Avail cDNA.
XX
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KW ss.
XX
OS Hepatitis C virus.
XX
Key Location/Qualifiers
FH 1801..7758
FT /*tag= a
FT /*product= "HCVreplBartman polypeptide"
FT misc_feature 7766
FT /*tag= b
FT /*note= "Nucleotide creating Avail site"
XX
PN WO200189364-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US016822.
XX
PR 23-MAY-2000; 2000US-00576989.
XX
PA (UNIT) UNIV WASHINGTON.
XX
PI Rice CM, Blight KJ;
PI MPI; 2002-066755/09.
PI P-PSDB; AAB51717.
XX
PT Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
XX
PS Claim 44; Page 69-71; 174pp; English.
XX
CC The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV

```

```

CC protein and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) replBartman/Avail cDNA
XX
SQ Sequence 7989 BP; 1647 A; 2369 C; 2242 G; 1731 T; 0 U; 0 Other;
Query Match 85.5%; Score 884.6; DB 6; Length 7989;
Best Local Similarity 90.9%; Pred. No. 1.2e-227;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 1 ACAGAGTGTGATGAGGAGTGGCGCTGACAGAGTACGCTCCGCGCTGCAACCTCTCTACCG 60
DB 5092 ACAGAGTGTGATGAGGAGTGGCGCTGACAGAGTACGCTCCGCGCTGCAACCTCTCTACCG 5151
QY 61 GAGGAGTGTGATGAGGAGTGGCGCTGACAGAGTACGCTCCGCGCTGCAACCTCTCTACCG 120
DB 5152 GAGGAGTGTGATGAGGAGTGGCGCTGACAGAGTACGCTCCGCGCTGCAACCTCTCTACCG 5211
QY 121 GAGCGCGAAGCGGATGTATGATGATGATGATGATGATGATGATGATGATGATGATG 180
DB 5212 GAGCGCGAAGCGGATGTATGATGATGATGATGATGATGATGATGATGATGATGATG 5271
QY 181 GCAGAGACGAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 5272 GCAGAGACGAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5331
QY 241 GCTAGCAGTGTGTGCGCTTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 5332 GCTAGCAGTGTGTGCGCTTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5391
QY 301 GACGCTGACCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 5392 GACGCTGACCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5451
QY 361 CGCGTGAAGTCAAGAGCAAGGATGATCTGACTCTTTGACCGCGCTCGAGGAGGAG 420
DB 5452 CGCGTGAAGTCAAGAGCAAGGATGATCTGACTCTTTGACCGCGCTCGAGGAGGAGGAG 5511
QY 421 GAGGATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 5512 GAGGATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5571
QY 481 CCAAGCATGCGCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 5572 CCAAGCATGCGCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5631
QY 541 GCCCGGAGTCAAGTCCCTCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 600
DB 5632 GCCCGGAGTCAAGTCCCTCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 5691
QY 601 CCTATACACCTCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 5692 CCTATACACCTCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5751
QY 661 GCCCTGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 5752 GCCCTGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5811
QY 721 GGCAGCGCAAGCGCGCTCTGAGCAATCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 780

```

Db 5812 GGCACGGACGCGCTCTCTCTGACGACGCCCTCCGACGACGCGGAGTCCGACGTT 5871  
Qy 781 GAGTCGATTCCTCCATGCGCCCTTGAAGGGAGAGCCGGAGACCCGATCTCAGCGAC 840  
Db 5872 GAGTCGATTCCTCCATGCGCCCTTGAAGGGAGAGCCGGAGATCCGATCTCAGCGAC 5931  
Qy 841 GGGTCCTGCTACCGCTGAGTGAAGAGCGCGGTGAGAGAGCTGCTGCTCGATGTC 900  
Db 5932 GGGTCCTGCTACCGCTGAGTGAAGAGCGCTGCTGCTGCTCGATGTC 5991  
Qy 901 TACACATGACGAGCGCTGATCAACGACGCTGCGAGAGAAAGCAAGTCCGACATC 960  
Db 5992 TACACATGACGAGCGCTGATCAACGACGCTGCGAGAGAAAGCAAGTCCGACATC 6051  
Qy 961 AACGCGTTGAGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Db 6052 AATGCACTGAGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6111  
Qy 1021 AGCGCAAGCCGCGG 1035  
Db 6112 AGCGCAAGCCGCGG 6126

## RESULT 12

ADJ57845 standard; DNA; 7989 BP.

ADJ57845;

06-MAY-2004 (first entry)

HCV replicon encoding sequence.

hepatitis C virus; HCV; Antiinflammatory; Hepatotropic; Virucide; ds;  
HCV replicon.

Unidentified.

Key Location/Qualifiers

FT CDS 1801..7759

FT /tag= a

FT /product= "HCV replicon"

PN WO2004015131-A2.

XX 19-FEB-2004.

XX 12-AUG-2003; 2003WO-US025260.

XX 12-AUG-2002; 2002US-0402661P.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Gao M, Lemm JA, O'Boyle DR, Nower P, Rigat K, Sun J;

XX WPI; 2004-180685/17.

XX P-PSDB; ADJ57846.

XX Use of hepatitis C virus assays or reporter assays, e.g. identifying a

XX compound that modulates the activity of a gene of interest.

XX Claim 3; SEQ ID NO 1; 45bp; English.

XX The present invention relates to the use of hepatitis C virus (HCV)  
XX assays for identifying a compound that inhibits HCV RNA replication and  
XX reporter assays for identifying a compound that modulates the activity of  
XX a gene of interest. The assays are useful for identifying a compound that  
XX inhibits HCV RNA replication or for identifying a compound that modulates  
XX the activity of a gene of interest. The HCV assay is useful for high  
XX throughput screening that quantifies both the amount of HCV RNA  
XX replication inhibitory activity associated with a test compound and the  
XX amount of cytotoxicity associated with the test compound. The compound is

CC useful for treating hepatitis C infection. Assays of the invention have  
CC distinct advantages when compared to qRT-PCR or other methods in that  
CC assays of the invention may take place in situ in a detergent based crude  
CC cell lysate, which requires no further preparation prior to performing  
CC the assays. The assays do not also involve numerous manipulations to add  
CC or subtract reagents after addition of test compounds and are desirably  
CC based on a viral protein which is required by the HCV replicon for  
CC replication. The present sequence represents a HCV replicon encoding  
CC sequence used in the assay of the invention.

SQ Sequence 7989 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other;

Query Match 85.5%; Score 884.6; DB 12; Length 7989;

Best Local Similarity 90.9%; Pred. No. 1.2e-227;

Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 1 ACAAGATGATGGGGTGGGCTGACAGGTAAGCTCCGGGTGCAAACTCTCTTACGG 60  
Db 5092 ACAAGATGATGGGGTGGGCTGACAGGTAAGCTCCGGGTGCAAACTCTCTTACGG 5151  
Qy 61 GAGAGGTCAATTCAGGTCGAGGCTCAACCAATACCTGTTGGGTGAGCTCCATGC 120  
Db 5152 GAGAGGTCAATTCAGGTCGAGGCTCAATTAATCTGTTGGGTGAGCTCCATGC 5211  
Qy 121 GAGCCGGAACCGAGTGTAGAGTGTCACTTCAATGCTACGACCCCTCCACATCA 180  
Db 5212 GAGCCGGAACCGAGTGTAGAGTGTCACTTCAATGCTACGACCCCTCCACATTA 5271  
Qy 181 GAGAGAGGCTTAAGCGGAGGCTGCGCAGGGGGTCTCCCTCTCTTGGCCAGCTTCA 240  
Db 5272 GCGGAGAGGCTTAAGCGGAGGCTGCGCAGGGGGTCTCCCTCTTGGCCAGCTCA 5331  
Qy 241 GCTAGCCAGTTGTCTGCGCCCTTCTCCGAGGCGCATATCAATACCAAAATGACTTCCA 300  
Db 5332 GCTAGCCAGTTGTCTGCGCCCTTCTTGAAGGCAATGACTATCCGTAATCTCCCG 5391  
Qy 301 GACGCTGACCTATCGAGGCAACCTCTGTGCGGAGTGAAGTGGGCGGAGCATTA 360  
Db 5392 GACGCTGACCTATCGAGGCAACCTCTGTGCGGAGTGAAGTGGGCGGAGCATCA 5451  
Qy 361 GCGGTGAGTCAAGAGAAACAAGTATTAATCTTGAATCTTTCGACCCGCTCCGAGGAG 420  
Db 5452 GCGGTGAGTCAAGAGAAACAAGTATTAATCTTTCGAGCGCTCCGAGGAG 5511  
Qy 421 GAGATGAGCGGAGAGTGTCCGTCGCGGAGAGATCTCGCGAAATCCAGAAATTC 480  
Db 5512 GAGATGAGAGAGAGTGTCCGTCGCGGAGAGATCTCGCGAGTCCAGAAATTC 5571  
Qy 481 CCAGGATGCCCGATGGGACGCGCGGATTACAACCTCTGCTGAGTCTCTGAGG 540  
Db 5572 CCAGGATGCCCGATGGGACGCGCGGATTACAACCTCTGAGTCTCTGAGG 5631  
Qy 541 GCCCGGATCAATGTCCTTCCAGTGTGATAGTGGTCCCACTGCCAATCTTAAGCCCT 600  
Db 5632 GACCGGATCAATGTCCTTCCAGTGTGATAGTGGTCCCACTGCCAATCTTAAGCCCT 5691  
Qy 601 CCTATACCACTCCACGAGAGAGAGCACTTGTCTGACAGATCCACCGTCTCTCT 660  
Db 5692 CGATACCACTCCACGAGAGAGAGCACTTGTCTGATAGATCTACCGTCTCTCT 5751  
Qy 661 GCCCTGGAGAGCTTGCACAAAGCTTTGTGATCTCGGACCGTGGCGCTGACAGC 720  
Db 5752 GCTTGGGAGAGCTTGCACAAAGCTTGCAGAGCTCCGATGCTGGCGCTGACAGC 5811  
Qy 721 GGCACGGACACCGCCCTCTCTGACCAATCTTCCGACGCGGAGAGATCTGACGTT 780  
Db 5812 GGCACGGACACCGCCCTCTCTGACCAATCTTCCGACGCGGAGATCTGACGTT 5871  
Qy 781 GAGTCGATTCCTCCATGCGCCCTTGAAGGGAGAGCCGGAGATCCGATCTCAGCGAC 840  
Db 5872 GAGTCGATTCCTCCATGCGCCCTTGAAGGGAGAGCCGGAGATCCGATCTCAGCGAC 5931  
Qy 841 GGGTCCTGCTACCGCTGAGTGAAGAGCGGTGAGAGAGCTGCTGCTCGATGTC 900





AA147276  
ID AA147276 standard; DNA; 7992 BP.  
AC AA147276;  
DT 30-AUG-2002 (first entry)  
XX Hepatitis C virus sub-genomic replicon clone I377-NS3-3'UTR.  
DE Hepatitis C virus sub-genomic replicon clone I377-NS3-3'UTR.  
XX Hepatitis; HCV, core-neo; NS3 proteinase/helicase; vaccine; diagnosis;  
KM virucide; hepatotropic; gene therapy; anti-viral; gene; ds.  
XX Hepatitis C virus.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 342..1181  
FT /tag= a  
FT /product= "core-neo fusion protein"  
FT 1801..7758  
FT /tag= b  
FT /product= "NS3 proteinase/helicase"  
XX  
XX WO200238793-A2.  
XX  
XX 16-MAY-2002.  
XX  
XX  
PF 02-NOV-2001; 2001WO-US046350.  
XX  
XX 07-NOV-2000; 2000US-0245866P.  
XX  
XX (ANAD-) ANADYS PHARM INC.  
XX  
XX Bichko V;  
XX  
XX WPI; 2002-490082/52.  
DR P-PSDB; AA018000; AA018001.  
XX  
XX Novel nucleic acid encoding replication competent recombinant hepatitis C  
PT virus genome useful for screening anti-hepatitis C virus therapeutics and  
PT for vaccine development.  
XX  
XX  
XX Claim 6; Page 43-47; 85pp; English.  
XX  
XX The present invention provides protein and coding sequences from  
CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and  
CC able to replicate efficiently when transfected into a susceptible cell  
CC line without reducing the growth rate of the cell line by more than 10  
CC fold. The sequences are useful for screening for anti-HCV therapeutics,  
CC for detecting antibodies to HCV in a biological sample such as blood,  
CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,  
CC for deriving authentic HCV components such as replication-competent non-  
CC infectious, replication-defective infection-component, and replication-  
CC defective non-infectious HCV, in gene therapy or gene vaccination  
CC targeted to hepatic tissue for treating an animal infected or susceptible  
CC to HCV infection and for studying HCV infection and propagation. The  
CC present sequence is a clone of a fragment of the HCV genome which encodes  
CC the core-neo and NS3 proteinase/helicase proteins  
XX  
XX Sequence 7992 BP; 1648 A; 2369 C; 2243 G; 1732 T; 0 U; 0 Other;  
SQ

Query Match 85.5%; Score 884.6; DB 6; Length 7992;  
Best local similarity 90.9%; Pred. No. 1.2e-227;  
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1 ACAGAGTGGATGGGTCGGCTGCAAGTACGCTCCGCGTGAACCTCTCTCAACG 60  
DB 5092 ACAGAGTGGATGGGTCGGCTGCAAGTACGCTCCGCGTGAACCTCTCTCAACG 5151  
QY 61 GAGGAGTCACTTCCAGGTCGGGCTCAACCAATACCTGTTGGTGGAGCTCCCATCC 120  
DB 5152 GAGGAGTCACTTCCAGGTCGGGCTCAACCAATACCTGTTGGTGGAGCTCCCATCC 5211  
QY 121 GAGCCGAACCGGATGAGAGTGTCACTTCCATGCTCAACGACCCCTCCCATCAACA 180

DB 5212 GAGCCGAACCGGATGAGAGTGTCACTTCCATGCTCAACGACCCCTCCCATCAACA 5271  
QY 181 GAGGAGTCACTTCCAGGTCGGGCTCAACCAATACCTGTTGGTGGAGCTCCCATCC 240  
DB 5272 GAGGAGTCACTTCCAGGTCGGGCTCAACCAATACCTGTTGGTGGAGCTCCCATCC 5331  
QY 241 GCTAGCCAGTGTCTGCGCCCTCTCCGAGGCGCATATCATTAACCAATAGACTCCCA 300  
DB 5332 GCTAGCCAGTGTCTGCGCCCTCTCCGAGGCGCATATCATTAACCAATAGACTCCCA 5391  
QY 301 GACGCTGACCTCATGAGAGGCAACCTCTGTGGCGGATGAGATGGCGGGAACATTAAC 360  
DB 5392 GACGCTGACCTCATGAGAGGCAACCTCTGTGGCGGATGAGATGGCGGGAACATTAAC 5451  
QY 361 CGCGTGGATGAGAGGCAACCAAGTATTAATCTGATCTTTTCCAGCCGCTCCGAGCGAG 420  
DB 5452 CGCGTGGATGAGAGGCAACCAAGTATTAATCTTGTGAGCTTTTCCAGCCGCTCCAGCGAG 5511  
QY 421 GAGGATGAGCGGGAAGTGTCCGTCGCCGCGGAGATCTCGCGGAATCCCAAGAAATTCCTCA 480  
DB 5512 GAGGATGAGCGGGAAGTGTCCGTCGCCGCGGAGATCTCGCGGAAGTCCAGGAATTCCTCT 5571  
QY 481 CCAGGATGCCCGGATGGGCAAGCCCGGATTAACACCTCTCGCTGTGAGTCTTGAAAG 540  
DB 5572 CCAGGATGCCCGGATGGGCAAGCCCGGATTAACACCTCTCGCTGTGAGTCTTGAAAG 5631  
QY 541 GCCCGGATCTAGTCCCTCTCCAGTGTATATGGTGTCCCACTCCCACTTAAGACCCCT 600  
DB 5632 GACCCGGACTAGTCCCTCTCCAGTGTATATGGTGTCCCACTCCCACTTAAGACCCCT 5691  
QY 601 CCTATACCACTCCAGGAGAAAGAGAGCAGTTGTCTGACAGATCCACCGTGTCTTCT 660  
DB 5692 CGATACCACTCCAGGAGAAAGAGAGCAGTTGTCTGACAGATCCACCGTGTCTTCT 5751  
QY 661 GCCCTGGGAGCTTGGCCCAAAAGCTTTTGTAGCTTCGGAACCGTGGCGGTGCAGAC 720  
DB 5752 GCCCTGGGAGCTTGGCCCAAAAGCTTTTGTAGCTTCGGAACCGTGGCGGTGCAGAC 5811  
QY 721 GGCACGGAGACCGCCCTCTCTGACCAATCTTCCAGCAGCGCGAGAGATCTGACGTT 780  
DB 5812 GGCACGGAGACCGCCCTCTCTGACCAATCTTCCAGCAGCGCGAGATCTGACGTT 5871  
QY 781 GAGTGTATCTCTCATGACCCCTTTGAGGGGAGCGGGGAGACCCGATCTGACGAC 840  
DB 5872 GAGTGTATCTCTCATGACCCCTTTGAGGGGAGCGGGGAGATCCGATCTGACGAC 5931  
QY 841 GGGTCTTGTCTTACCGTGAAGTGAAGAGCCGATGAGACGTCTGTGCTGATGTC 900  
DB 5932 GGGTCTTGTCTTACCGTGAAGTGAAGAGCCGATGAGACGTCTGTGCTGATGTC 5991  
QY 901 TACACATGAGAGCGGCTCTGTATCAAGCATGCGCTGGAGGAAAGAGACTGCCATC 960  
DB 5992 TACACATGAGAGCGGCTCTGTATCAAGCATGCGCTGGAGGAAAGAGACTGCCATC 6051  
QY 961 AACGGTGAAGCACTCTTGTGTGTCACCAACAATGCTACGATCAACATCCGCG 1020  
DB 6052 AATGCACTGAAGCACTTGTGTGTCACCAACAATGCTACGATCAACATCTTGC 6111  
QY 1021 AGCGCAACCGAGCG 1035  
DB 6112 AGCGCAACCGAGCG 6126

RESULT 15  
AA147280  
ID AA147280 standard; DNA; 7992 BP.  
AC AA147280;  
DT 30-AUG-2002 (first entry)  
XX Hepatitis C virus sub-genomic replicon recombinant clone HCVR22.



XX Hepatitis; HCV, core-neo; NS3 proteinase/helicase; vaccine; diagnosis;  
 KM virucide; hepatotropic; gene therapy; anti-viral; gene; ds.  
 XX Hepatitis C virus.  
 OS  
 PN MO200238793-A2.  
 XX  
 PD 16-MAY-2002.  
 XX  
 PF 02-NOV-2001; 2001WO-US046350.  
 XX  
 PR 07-NOV-2000; 2000US-0245866P.  
 XX  
 PA (ANAD-) ANADYS PHARM INC.  
 XX  
 PI Bichko V;  
 XX  
 DR WPI; 2002-490082/52.  
 XX  
 PT Novel nucleic acid encoding replication competent recombinant hepatitis C  
 PT virus genome useful for screening anti-hepatitis C virus therapeutics and  
 PT for vaccine development.  
 XX  
 PS Claim 10; Page 66-70; 85pp; English.  
 XX  
 CC The present invention provides protein and coding sequences from  
 CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and  
 CC able to replicate efficiently when transfected into a susceptible cell  
 CC line without reducing the growth rate of the cell line by more than 10  
 CC fold. The sequences are useful for screening for anti-HCV therapeutics,  
 CC for detecting antibodies to HCV in a biological sample such as blood,  
 CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,  
 CC for deriving authentic HCV components such as replication-competent non-  
 CC infectious, replication-defective infection-component, and replication-  
 CC defective non-infectious HCV, in gene therapy or gene vaccination  
 CC targeted to hepatic tissue for treating an animal infected or susceptible  
 CC to HCV infection and for studying HCV infection and propagation. The  
 CC present sequence is a clone of a fragment of the HCV genome designated  
 CC HCVR2  
 CC  
 SQ Sequence 7992 BP; 1646 A; 2368 C; 2245 G; 1733 T; 0 U; 0 Other;

Query Match 85.5%; Score 884.6; DB 6; Length 7992;  
 Best Local Similarity 90.9%; Pred. No. 1.2e-227;  
 Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1 ACAGAGTGGATGGGCTGCGGCTGACAGGTACGCTCCGCGGAGCAAACTCTCTACCG 60  
 DB 5092 ACAGAGTGGATGGGCTGCGGCTGACAGGTACGCTCCGCGGAGCAAACTCTCTACCG 5151  
 QY 61 GAGGAGTGCATTCAGGTCGGGCTCAACCAATACCTGTTGGTGCAGCTCCATGC 120  
 DB 5152 GAGGAGTGCATTCAGGTCGGGCTCAATCAATACCTGTTGGTGCAGCTCCATGC 5211  
 QY 121 GAGCCCGAACCAGATATGACAGTCTCATCTGCTACGACCGACCCCTCCACATCA 180  
 DB 5212 GAGCCCGAACCAGATATGACAGTCTCATCTGCTACGACCGACCCCTCCACATCA 5271  
 QY 181 GCAGAGACGGCTAAGCGCAGGCTGCGCAGGAGGCTCCGCCCTCTTGGCAGCTCTTCA 240  
 DB 5272 GCAGAGACGGCTAAGCGCAGGCTGCGCAGGAGGCTCCGCCCTCTTGGCAGCTCTTCA 5331  
 QY 241 GCTAGCAGTGTCTGCGCCCTTCTCTGAAAGCGACATTAACCAAAATGACTTCCA 300  
 DB 5332 GCTAGCAGTGTCTGCGCCCTTCTCTGAAAGCGACATTAACCAAAATGACTTCCA 5391  
 QY 301 GACGCTGACCTATCGAGCGCAACCTCTGTCGGCGGATGAGATGGCGGGAATTACC 360  
 DB 5392 GACGCTGACCTATCGAGCGCAACCTCTGTCGGCGGATGAGATGGCGGGAATTACC 5451  
 QY 361 CGGCTGAGTCAAGAAACAAGTAGTAATCTGGAATCTTTCGACCCGCTCCGAGCGGAG 420

DB 5452 CGCGTGAAGTCAAAAAATAGGTAGTAATTTTGAAGCTTTTCGAGCCGCTCCAGCGGAG 5511  
 QY 421 GAGGATGACGGGAAAGTGTCCGTCCCGGAGATCTCTCGGAAATCCAGAAATTTCCA 480  
 DB 5512 GAGGATGAGGAAAGTATCCGTTCCGGGAGATCTCTCGGAGGTCGAGAAATTCCT 5571  
 QY 481 CCAGCGATGCCCGATGGGACGCGCGGATTAACAACCTCCGCTGTGAGTCTTGAAG 540  
 DB 5572 CGAGCATGCCCATATGGGACGCGCGGATTAACAACCTCTCACTTTAGTCTTGAAG 5631  
 QY 541 GCCCGGACATAGTCTCCATGATGATACATGGGATGCGCATGCACTTAAGACCCCT 600  
 DB 5632 GACCGGACATAGTCTCCATGATGATACATGGGATGCGCATGCACTTAAGACCCCT 5691  
 QY 601 CCTATTCACCTTCACGAGAAAGAGACAGTTGTTCTACAGATTCACCGTGTCTTCT 660  
 DB 5692 CCGATACCACTTCACGAGAAAGAGACAGTTGTTCTACAGATTCACCGTGTCTTCT 5751  
 QY 661 GCCCTGCGGAGCTTGGCAAAAGGCTTTGATAGTCCGACCGTCCGCTCCAGC 720  
 DB 5752 GCTTGGCGGAGCTTGGCAAAAGGCTTTGATAGTCCGACCGTCCGCTCCAGC 5811  
 QY 721 GGCACGCGCAACCGCCCTCTGACCAATCTCTCGACGAGCGGACAGATCTGACGTT 780  
 DB 5812 GGCACGCGCAACCGCCCTCTGACCAATCTCTCGACGAGCGGACAGATCTGACGTT 5871  
 QY 781 GAGTGTATTTCTCCATGCCCCCTTTGAGGGGAGCGGGGACCCGATCTCAGCGAC 840  
 DB 5872 GAGTGTATTTCTCCATGCCCCCTTTGAGGGGAGCGGGGACCCGATCTCAGCGAC 5931  
 QY 841 GGGTCTGCTTACCGTGAAGAGGAGCGGGGAGGAGTCTGCTGCTGATGATGTC 900  
 DB 5932 GGGTCTGCTTACCGTGAAGAGGAGCGGGGAGGAGTCTGCTGCTGATGATGTC 5991  
 QY 901 TACACATGACAGGCGCTCTGATCAAGCCATGCTGCGAGAGAAAGCAAGCTGCCATC 960  
 DB 5992 TACACATGACAGGCGCTCTGATCAAGCCATGCTGCGAGAGAAAGCAAGCTGCCATC 6051  
 QY 961 AACGCTTGAAGCACTTTTGTGCTGATCAACCAATGTTAGCTTACCATTCGCG 1020  
 DB 6052 AACGCTTGAAGCACTTTTGTGCTGATCAACCAATGTTAGCTTACCATTCGCG 6111  
 QY 1021 AGCGCAAGCCAGCGG 1035  
 DB 6112 AGCGCAAGCTGCGG 6126

Search completed: February 19, 2005, 01:26:44  
 Job time : 582.412 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2005, 00:21:46 ; Search time 178.108 Seconds  
(without alignments)  
9508.540 Million cell updates/sec

Title: US-09-664-363-4

Perfect score: 1035  
Sequence: 1 ACAGAGTCGATGCGGTGCG.....CCGCGAGCCGACGCGCG 1035

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1035	100.0	1035	US-08-191-160-4	Sequence 4, Appl
2	1025.4	99.1	3750	US-08-191-160-22	Sequence 22, Appl
3	885.8	85.6	8642	US-10-029-907-2	Sequence 2, Appl
4	884.6	85.5	7989	US-09-539-601-10	Sequence 10, Appl
5	884.6	85.5	8001	US-09-539-601-7	Sequence 7, Appl
6	884.6	85.5	8637	US-09-539-601-4	Sequence 4, Appl
7	884.6	85.5	8638	US-10-029-907-6	Sequence 6, Appl
8	884.6	85.5	8638	US-10-029-907-24	Sequence 24, Appl
9	884.6	85.5	8639	US-10-029-907-1	Sequence 1, Appl
10	884.6	85.5	8643	US-09-539-601-13	Sequence 13, Appl
11	884.6	85.5	11076	US-09-539-601-1	Sequence 1, Appl
12	883	85.3	8638	US-10-029-907-7	Sequence 7, Appl
13	883	85.3	8638	US-10-029-907-25	Sequence 25, Appl
14	883	85.3	8643	US-10-029-907-4	Sequence 4, Appl
15	883	85.3	8648	US-10-029-907-5	Sequence 5, Appl
16	879.8	85.0	8001	US-09-539-601-16	Sequence 16, Appl
17	879.8	85.0	8001	US-09-539-601-22	Sequence 22, Appl
18	879.8	85.0	11076	US-09-539-601-19	Sequence 19, Appl
19	879.8	85.0	11076	US-09-539-601-25	Sequence 25, Appl
20	878.2	84.9	1176	US-08-537-811-41	Sequence 41, Appl
21	878.2	84.9	9472	US-08-150-204E-96	Sequence 96, Appl
22	876.6	84.7	8001	US-09-539-601-28	Sequence 28, Appl
23	876.6	84.7	11076	US-09-539-601-31	Sequence 31, Appl
24	872.4	84.3	2991	US-08-324-977-49	Sequence 49, Appl
25	872.4	84.3	2991	US-08-384-616-49	Sequence 49, Appl
26	872.4	84.3	2991	US-08-904-686A-49	Sequence 49, Appl
27	872.4	84.3	2991	US-09-315-850-49	Sequence 49, Appl

28	872.4	84.3	7863	US-08-324-977-35	Sequence 35, Appl
29	872.4	84.3	7863	US-08-384-616-35	Sequence 35, Appl
30	872.4	84.3	7863	US-08-904-686A-35	Sequence 35, Appl
31	872.4	84.3	7863	US-09-315-850-35	Sequence 35, Appl
32	872.4	84.3	7917	US-08-324-977-31	Sequence 31, Appl
33	872.4	84.3	7917	US-08-384-616-31	Sequence 31, Appl
34	872.4	84.3	7917	US-08-904-686A-31	Sequence 31, Appl
35	872.4	84.3	7917	US-09-315-850-31	Sequence 31, Appl
36	872.4	84.3	9030	US-08-324-977-13	Sequence 13, Appl
37	872.4	84.3	9030	US-08-384-616-13	Sequence 13, Appl
38	872.4	84.3	9030	US-08-904-686A-13	Sequence 13, Appl
39	872.4	84.3	9030	US-09-315-850-13	Sequence 13, Appl
40	872.4	84.3	9416	US-08-324-977-1	Sequence 1, Appl
41	872.4	84.3	9416	US-08-384-616-1	Sequence 1, Appl
42	872.4	84.3	9416	US-08-904-686A-1	Sequence 1, Appl
43	872.4	84.3	9416	US-09-315-850-1	Sequence 1, Appl
44	872.4	84.3	9416	US-08-823-895A-27	Sequence 27, Appl
45	862.2	83.3	9595	US-09-014-416-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-08-191-160-4  
Sequence 4, Application US/08191160  
Patent No. 6210675  
GENERAL INFORMATION:  
APPLICANT: Highfield, Peter Edmund  
APPLICANT: Rodgers, Brian Colin  
APPLICANT: Tedder, Richard Seton  
APPLICANT: Barbara, John Anthony James  
TITLE OF INVENTION: Viral Agent  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Rothwell, Pigg, Ernst & Kurz  
STREET: 1700 K Street  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage  
COMPUTER: IBM AT compatible  
OPERATING SYSTEM: MS-DOS V3.2  
SOFTWARE: Wordperfect 5.0 (DOS text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/191,160  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/628,516  
FILING DATE: 17 DEC 1990  
APPLICATION NUMBER: UK 89 28 562.1  
FILING DATE: 18 DEC 1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 90 04 414.0  
FILING DATE: 27 FEB 1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 90 04 814.1  
FILING DATE: 03 MAR 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: E. Anthony Figg  
REGISTRATION NUMBER: 27,195  
REFERENCE/DOCKET NUMBER: 1645-103A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 833-5740  
TELEFAX: (202) 833-5744  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1035 base pairs  
TYPE: nucleotide with corresponding protein  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: cDNA to genomic RNA  
ORIGINAL SOURCE: human; serum infectious for PT-NANBH  
IMMEDIATE SOURCE: clone JG3 from cDNA library in lambda gt11  
FEATURE:  
LOCATION: from 1 to 1035 bp portion of the PT-NANBH  
LOCATION: polyprotein  
OTHER INFORMATION: probably encodes viral non-structural  
OTHER INFORMATION: proteins  
US-08-191-160-4

Query Match 100.0%; Score 1035; DB 3; Length 1035;  
Best Local Similarity 100.0%; Pred. No. 4,3e-276;  
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ACAGAAAGTGAATGGGGTGGCGGTGACAGAGTACGCTCCGCGCTGCAAACTCTCTACGG 60
DB 1 ACAGAAAGTGAATGGGGTGGCGGTGACAGAGTACGCTCCGCGCTGCAAACTCTCTACGG 60
QY 61 GAGGAGGTACATTCCAGGTCCGGGCTCAACCAATACCTGGTGGGTGGAGTCCCATGC 120
DB 61 GAGGAGGTACATTCCAGGTCCGGGCTCAACCAATACCTGGTGGGTGGAGTCCCATGC 120
QY 121 GAGCCCGAACCAGGATGACAGTGTCACTTCCATGCTCAGCAGCCCTCCACATCACA 180
DB 121 GAGCCCGAACCAGGATGACAGTGTCACTTCCATGCTCAGCAGCCCTCCACATCACA 180
QY 181 GCAGAGACGGCTAAGCGCAGCGGTGCGCAGGGGGTCTCCCTCCTTGGCCAGCTTTCA 240
DB 181 GCAGAGACGGCTAAGCGCAGCGGTGCGCAGGGGGTCTCCCTCCTTGGCCAGCTTTCA 240
QY 241 GCTAGCCAGTGTCTGGCCCTTCTCGAAGGGGACATTAACCAATGACTTCCCA 300
DB 241 GCTAGCCAGTGTCTGGCCCTTCTCGAAGGGGACATTAACCAATGACTTCCCA 300
QY 301 GAGCTGACCTCATCGAGGCCAACCCTCTGTGGCGGATGAGTGGGGACATTACC 360
DB 301 GAGCTGACCTCATCGAGGCCAACCCTCTGTGGCGGATGAGTGGGGACATTACC 360
QY 361 CGGTGAGTCAAGAGAACAGTGAATCTTGTGACTCTTTGACCCGCTCCGAGCGAG 420
DB 361 CGGTGAGTCAAGAGAACAGTGAATCTTGTGACTCTTTGACCCGCTCCGAGCGAG 420
QY 421 GAGGATGACGGGGAATGTCCTCCGCGGGGAGATCTCGGAAATCCAAATTTCCA 480
DB 421 GAGGATGACGGGGAATGTCCTCCGCGGGGAGATCTCGGAAATCCAAATTTCCA 480
QY 481 CCAGCGATCCCGCATGGGACAGCCCGGATTAACAACCTCCGCTGCTGAGTCTGGAAG 540
DB 481 CCAGCGATCCCGCATGGGACAGCCCGGATTAACAACCTCCGCTGCTGAGTCTGGAAG 540
QY 541 GCCCGGACTAGTCTCCCTCAAGTGTACATGGGTGCCACTGCGACTTAAGACCCCT 600
DB 541 GCCCGGACTAGTCTCCCTCAAGTGTACATGGGTGCCACTGCGACTTAAGACCCCT 600
QY 601 CCTATACCACTCCAGGAAAGAGAGTGTCTGAAGAAATCCACCGTCTTCT 660
DB 601 CCTATACCACTCCAGGAAAGAGAGTGTCTGAAGAAATCCACCGTCTTCT 660
QY 661 GCCCTGGCGGAGTTTCCCAAAAGTTTGTGAGTCCGAGCCGTCGCGTCCAGCAGC 720
DB 661 GCCCTGGCGGAGTTTCCCAAAAGTTTGTGAGTCCGAGCCGTCGCGTCCAGCAGC 720
QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGAGCGGAGAGATCTGAGCTT 780
DB 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGAGCGGAGAGATCTGAGCTT 780
QY 781 GAGTGTATTCCTCAATGCCCCCTTGAAGGGGAGCCGGGGAGCCCGATCTCAGCGAC 840
DB 781 GAGTGTATTCCTCAATGCCCCCTTGAAGGGGAGCCGGGGAGCCCGATCTCAGCGAC 840
```

```
QY 841 GGGTCTGTCTACCGTAGTAGAGAGCCGGTAGAGAGAGTCTGCTGCTCGATGTC 900
DB 841 GGGTCTGTCTACCGTAGTAGAGAGCCGGTAGAGAGAGTCTGCTGCTCGATGTC 900
QY 901 TACACATGACAGCGGCTCTGATCAGCCATGCGCTGGGAGGAAAGCAAGTGGCCATC 960
DB 901 TACACATGACAGCGGCTCTGATCAGCCATGCGCTGGGAGGAAAGCAAGTGGCCATC 960
QY 961 AACGGTTGAGCACTTTTGTGGGTGACCAACAATGTCCTAGTACCAATCCCGC 1020
DB 961 AACGGTTGAGCACTTTTGTGGGTGACCAACAATGTCCTAGTACCAATCCCGC 1020
QY 1021 AGCGCAAGCCAGCGG 1035
DB 1021 AGCGCAAGCCAGCGG 1035
```

```
RESULT 2
US-08-191-160-22
Sequence 22, Application US/08191160
Patent No. 6210675
GENERAL INFORMATION:
APPLICANT: Highfield, Peter Edmund
APPLICANT: Rodgers, Brian Colin
APPLICANT: Tedder, Richard Seton
APPLICANT: Barbara, John Anthony James
TITLE OF INVENTION: Viral Agent
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rothwell, Figg, Ernt & Kurz
STREET: 1700 K Street
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 KB storage
COMPUTER: IBM AT compatible
OPERATING SYSTEM: MS-DOS V3.2
SOFTWARE: Wordperfect 5.0 (DOS text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/191,160
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/628,516
FILING DATE: 17 DEC 1990
APPLICATION NUMBER: UK 89 28 562.1
FILING DATE: 18 DEC 1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 414.0
FILING DATE: 27 FEB 1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 814.1
FILING DATE: 03 MAR 1990
ATTORNEY/AGENT INFORMATION:
NAME: E. Anthony Figg
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 1645-103A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 833-5740
TELEFAX: (202) 833-5744
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3750 base pairs
TYPE: nucleotide with corresponding protein
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
ORIGINAL SOURCE:
ORGANISM: human; serum infectious for PT-NANBH
IMMEDIATE SOURCE:
LIBRARY: cDNA clones from 3' end of the genome
```

FEATURE:  
LOCATION: from 1 to 3750 bp portion of the PT-NANBH  
LOCATION: polypeptide  
OTHER INFORMATION: viral non-structural proteins  
US-08-191-160-22

Query Match 99.1%; Score 1025.4; DB 3; Length 3750;  
Best Local Similarity 99.4%; Pred. No. 3.1e-273;  
Matches 1029; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 ACAGAGTGAATGGAGTGGGCTGTCACAGTACGCTCCGCGTCAAACTCTCTACCG 60  
Db 1696 ACAGAGTGAATGGAGTGGGCTGTCACAGTACGCTCCGCGTCAAACTCTCTACCG 1755  
61 GAAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGTGGAGCTTCCATGC 120  
Db 1756 GAAGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGTGGAGCTTCCATGC 1815  
121 GAGCCGGAACGGATGATGAGTCACTTCATGCTACCGACCCCTCCCATCACA 180  
Db 1816 GAGCCGGAACGGATGATGAGTCACTTCATGCTACCGACCCCTCCCATCACA 1875  
181 GAGAGAGCGGCTAAGCGCAGAGTGGCCAGAGGAGTCTCCCTCTCTTGGCAGCTTCA 240  
Db 1876 GAGAGAGCGGCTAAGCGCAGAGTGGCCAGAGGAGTCTCCCTCTCTTGGCAGCTTCA 1935  
241 GCTAGCAGTGTCTGAGCCCTCTCTGAAAGGACATACCTTCAAAATGACTTCCA 300  
Db 1936 GCTAGCAGTGTCTGAGCCCTCTCTGAAAGGACATACCTTCAAAATGACTTCCA 1995  
301 GAGCGTGAATCAATGAGGCAACCTCTGAGCGGATGAGATGGGCGGCGGACATTAC 360  
Db 1996 GAGCGTGAATCAATGAGGCAACCTCTGAGCGGATGAGATGGGCGGCGGACATTAC 2055  
361 CGGCTGAGTCAAGAGAAAGAGTATCTCTGAGCTCTTCCAGCCGCTCCAGGCGAG 420  
Db 2056 CGGCTGAGTCAAGAGAAAGAGTATCTCTGAGCTCTTCCAGCCGCTCCAGGCGAG 2115  
421 GAGGATGAGCGGGAAGTGTCCGTCGCGCGGAGATCTCGGAAATCCAAATATCCCA 480  
Db 2116 GAGGATGAGCGGGAAGTGTCCGTCGCGCGGAGATCTCGGAAATCCAAATATCCCA 2175  
481 CCAAGCATCCCGCATGGGCAAGCCCGGATTAACCTCCGCTGAGTCTGAGAG 540  
Db 2176 CCAAGCATCCCGCATGGGCAAGCCCGGATTAACCTCCGCTGAGTCTGAGAG 2235  
541 GCGCCGGAATAGTCTCCCTCAGTGTATCATGGGTGCCATGCACTACTTAAGACCCCT 600  
Db 2236 GCGCCGGAATAGTCTCCCTCAGTGTATCATGGGTGCCATGCACTACTTAAGACCCCT 2295  
601 CCTATACCACTTCACGGAAGAAAGAGAGTGTCTGACAGAAATCCACCGTGTCTT 660  
Db 2296 CCTATACCACTTCACGGAAGAAAGAGAGTGTCTGACAGAAATCCACCGTGTCTT 2355  
661 GCGCTGCGGAGTGTCCCAAAAGGCTTTTGTAGTCTCCGAGACCTGCGCGTCAAGC 720  
Db 2356 GCGCTGCGGAGTGTCCCAAAAGGCTTTTGTAGTCTCCGAGACCTGCGCGTCAAGC 2415  
721 GCGCAGGCAACCGCTCTCTGACCAATCTCCGAGCAGCGCGGAGAGAGATCTGAGCT 780  
Db 2416 GCGCAGGCAACCGCTCTCTGACCAATCTCCGAGCAGCGCGGAGAGATCTGAGCT 2475  
781 GAGTGTATTTCTCTCAATGCCCCCTTTGAGGGGAGCGGGGAGACCCCGATCTCAAGGAC 840  
Db 2476 GAGTGTATTTCTCTCAATGCCCCCTTTGAGGGGAGCGGGGAGACCCCGATCTCAAGGAC 2535  
841 GGGTCTTGGTCAACGAGTGAAGAGGCGGAGAGAGTGTGAGTGTCTGAGTGTCC 900  
Db 2536 GGGTCTTGGTCAACGAGTGAAGAGGCGGAGAGAGTGTGAGTGTCTGAGTGTCC 2595  
901 TACACATGACAGCGCTGTATCAGCCCATGCGCTGCGAGAGAAAGCAAGTCCCATC 960  
Db 2596 TACACATGACAGCGCTGTATCAGCCCATGCGCTGCGAGAGAAAGCAAGTCCCATC 2655

961 AACGGTGAAGACATCTTGTGCTGACCAACAAGTGTCTAGCTACATCCCG 1020  
Db 2656 AACGGTGAAGACATCTTGTGCTGACCAACAAGTGTCTAGCTACATCCCG 2715  
1021 AGCGCAAGCCAGCGG 1035  
Db 2716 AGCGCAAGCCAGCGG 2730

RESULT 3  
US-10-029-907-2  
Sequence 2, Application US/10029907  
Patent No. 6706874

GENERAL INFORMATION:  
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
TITLE OF INVENTION: HEPATITIS C VIRUS  
FILE REFERENCE: 13/083  
CURRENT APPLICATION NUMBER: US/10/029,907  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/257,857  
PRIOR FILING DATE: 2000-12-22  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 8642

TYPE: DNA

ORGANISM: HCV

FEATURE:

NAME/KEY: CDS

LOCATION: (1802)... (8407)

NAME/KEY: variation

LOCATION: 6268

OTHER INFORMATION: r = a or g

NAME/KEY: variation

LOCATION: 4446

OTHER INFORMATION: r = a or g

US-10-029-907-2

Query Match 85.6%; Score 885.8; DB 4; Length 8642;  
Best Local Similarity 90.9%; Pred. No. 1.6e-234;  
Matches 941; Conservative 1; Mismatches 93; Indels 0; Gaps 0;

1 ACAGAGTGAATGGAGTGGGCTGTCACAGTACGCTCCGCGTCAAACTCTCTACCG 60  
Db 5741 ACAGAGTGAATGGAGTGGGCTGTCACAGTACGCTCCGCGTCAAACTCTCTACCG 5800  
61 GAGGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCATGC 120  
Db 5801 GAGGAGTCAATTCAGAGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCATGC 5860  
121 GAGCCGGAACGGATGATGAGTCACTTCATGCTCAACGACCCCTCCCATCACA 180  
Db 5861 GAGCCGGAACGGATGATGAGTCACTTCATGCTCAACGACCCCTCCCATCACA 5920  
181 GCGAGAGCGGCTAAGCGCAGAGTGGCCAGAGGAGTCTCCCTCTTGGCAGCTTCA 240  
Db 5921 GCGAGAGCGGCTAAGCGCAGAGTGGCCAGAGGAGTCTCCCTCTTGGCAGCTTCA 5980  
241 GCTAGCAGTGTCTGAGCCCTCTCTGAAAGGAGACATACCTTCAAAATGACTTCCA 300  
Db 5981 GCTAGCAGTGTCTGAGCCCTCTCTGAAAGGAGACATACCTTCAAAATGACTTCCA 6040  
301 GAGCTGACCTCATTCAGAGCAACCTCTGAGCGGATGAGATGGGCGGGAATATAC 360  
Db 6041 GAGCTGACCTCATTCAGAGCAACCTCTGAGCGGATGAGATGGGCGGGAATATAC 6100  
361 CGGCTGAGTCAAGAGAAAGAGTATCTCTGAGTCTTTTGAACCCGCTCCAGAGGAG 420  
Db 6101 CGGCTGAGTCAAGAGAAAGAGTATCTCTGAGTCTTTTGAACCCGCTCCAGAGGAG 6160  
421 GAGGATGAGCGGGAAGTGTCCGTCGCGCGGAGATCTCGGAAATCCAAATATCCCA 480

```

Db      6161 GAGGATGAGGAGGAGTATCCGTTCCGGCGAGATCTGTCGAGAGTCCAGGAAATTCCT 6220
OY      481  CCAGGATGCCCGCATGGGCAACGCCCGGATTACAACCTTCGGCTGCTGAAGTCTTGAAG 540
Db      6221 CGAGGATGCCCATATGGGCAACGCCCGGATTACAACCTTCAGTTGTGAATCTCTGGAG 6280
OY      541  GCGCCGGACTAGTCCCTCCAGTGGTACATGGGAGCCCACTGGCACTTACTAAGACCCCT 600
Db      6281 GACCCGACTAGTCTCTCCAGTGTACAGGAGTCCATTGGCCGCTCCGAAGCCCT 6340
OY      601  CCTATACCACCTCCACGAGAAAGAGACAGTTGTTCTGACAGATCCACCCTGTCTTCT 660
Db      6341 CCGATACCACTCCACGAGAAAGAGACGGTGTGCTGTAGATCTTACCTGCTTCT 6400
OY      661  GCCCTGGGGAGCTTGGCCACAAAGGCTTTTGGTAGCTCCGACCGTGGCCCTGACAGC 720
Db      6401 GCTTTGGGAGCTTCGCCAAGAAAGCTTGGCAGCTCCGAATGTCGGCCGTGACAGC 6460
OY      721  GGCACGGCAACCGCCCTCTGACCAATCTCCGACGACGCGGAGCAGATCTGACGT 780
Db      6461 GGCACGGCAAGGCTCTCTGACCAAGCTTCACAGCGCGACGAGCGGATCCGAGCTT 6520
OY      781  GAGTGTGATTTCTTCATATCCCCCTTTAGGGGAGCCCGGAGACCCGATCTACGAC 840
Db      6521 GAGTGTGATCTCTTCATATCCCCCTTTAGGGGAGCCCGGAGATCCGATCTACGAC 6580
OY      841  GGGTCTTGTCTACCGTGAAGTGAAGGACGGTGAAGACGTGATGCTGTGATGATGCC 900
Db      6581 GGGTCTTGTCTACCGTGAAGGAGGACTTGTGAAGACGTGATGCTGTGATGATGCC 6640
OY      901  TACACATGACAGGGGCTCTGATCAACGCATGCGGTGAGAGAAAGCAAGCTGCCATC 960
Db      6641 TACACATGACAGGGGCTCTGATCAACGCATGCGGTGAGAGAAAGCAAGCTGCCATC 6700
OY      961  AACGGCTGACCAACTCTTGTGGGTACACAGCAACAGTGTCTACGATCAACATCCGC 1020
Db      6701 AATGACTGACCAACTCTTGTGGGTACACAGCAACATGTTGTATGATCAACATCTGC 6760
OY      1021 AGCGCAAGCCAGCGG 1035
Db      6761 AGCGCAAGCCAGCGG 6775

RESULT 4
US-09-539-601-10
; Sequence 10, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartschlagel, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 7989
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct 1377/NS3-3'/wt
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1181)
; OTHER INFORMATION: hepatitis C virus core-neomycin phosphotransferase
; OTHER INFORMATION: fusion protein
; FEATURE:
; NAME/KEY: RBS

```

	LOCATION:	(1190)..(1800)	
	OTHER INFORMATION:	internal ribosome entry site from	
	OTHER INFORMATION:	encephalomyocarditis virus	
	FEATURE:		
	NAME/KEY:	CDS	
	LOCATION:	(1801)..(7758)	
	OTHER INFORMATION:	hepatitis C virus NS3 - 5B	
	FEATURE:		
	NAME/KEY:	3'UTR	
	LOCATION:	(7759)..(7989)	
	PUBLICATION INFORMATION:		
	AUTHORS:	Lohmann, Volker	
	AUTHORS:	Krner, Frank	
	AUTHORS:	Koch, Jan-Oliver	
	AUTHORS:	Herian, Ulrike	
	AUTHORS:	Theilmann, Lorenz	
	AUTHORS:	Bartenschlager, Ralf	
	TITLE:	Replication of subgenomic hepatitis C virus RNAs in a	
	TITLE:	hepatoma cell line	
	JOURNAL:	Science	
	VOLUME:	285	
	PAGES:	110-113	
	DATE:	1999-07-02	
	US-09-539-601-10		
Query Match	85.5%; Score 884.6; DB 4; Length 7989;		
Best Local Similarity	90.9%; Pred. No. 3.4e-234;		
Matches	941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;		
Dy	1 ACAGAGTGGATGGGCGTGCCGCTGCACAGGTACGCTCCGGCGTCGAACCTCTCTCAACG	60	
Dd	5092 ACAAAATGTGATGGGCGTGCCGCTTGCACAGGTACGCTCCAGGTGCAACCCTCCTACGG	5151	
Dy	61 GAGAGGTGCATTCCAGGTGGGCTCAACAATAACTGTTGGGTGCAGCTCCCATGC	120	
Dd	5152 GAGAGGTGCATTCTGTGTGGGCTCATTAATACCTGTTGGGTGCAGCTCCCATGC	52111	
Dy	121 GAGCCCAACCGATGTAGCAGTCTCATCTTCACGTCTCACGACCCCTCCCAATCACA	180	
Dd	5212 GAGCCCAACCGAGTAGCAGTCTCATCTTCATCTCACGCCCTCCCAATTAACG	52712	
Dy	181 GCAGAAGCGGTAGCCGACGCTGGCCAGGCGGCTCCCCCTCTTGGCCAGCTTCA	240	
Dd	5272 GCGAAGACGGTTAAGCGTAGGCTGGCCAGGAGATCTCCCCTCTTGGCCAGCTCA	53313	
Dy	241 GCTAGCCAGTTGTCTGGCCCTTCTCTCGAAGCGCATACATTACCAAATATGACTTCCA	300	
Dd	5332 GCTAGCCAGTGTCTGGCCCTTCTTGAAGCAATGACATTACCCGATGACTCCCG	5391	
Dy	301 GACGCTGACTCATCGAAGGCCAACCTCTGTGTGGCGCATGATGGCGGGACATTACC	360	
Dd	5392 GACGCTGACTCATCGAAGGCCAACCTCTGTGTGGCGCATGATGGCGGGACATTACC	5451	
Dy	361 CGCGTGGAGTCAGAGAACAGGTAGTATCTCTGACTCTTTGACCCGCTCCAGCGGAG	420	
Dd	5452 CGCGTGGAGTCAGAAAATTAGGTATGTTATTTGACTCTTTTCAGCCGCTCCAGCGGAG	55111	
Dy	421 GAGGATGAGCGGGAAGTGTCCGCTCCGCGCGAGATCTCTGCGAAATCCAGAAATTTCCA	480	
Dd	5512 GAGGATGAGAGGAAGTATCCGTTCCGCGCGAGATCTCTGCGAAATTTCCCT	55712	
Dy	481 CCAGCGATGCCCGGATGGGCAACGCCGCGAATTACAACCTCCGCTGGTGGATCTCTGAAG	540	
Dd	5572 CCAGCGATGCCCATATGGGCAACGCCGCGAATTACAACCTCCATCTTTAGATCTCTGGAAG	5631	
Dy	541 GCCCGGACTACGTCCCTCCAGTGTATCATGGGTGCCCATCTGACCACTTAAAGCCCT	600	
Dd	5632 GAACCGGACTACGTCCCTCCAGTGTATCATGGGTGCCCATCTGACCACTTAAAGCCCT	5691	
Dy	601 CCTATACCACTTCCAGGAGAAAGAGACAGTGTGTTGACAGAAATCCACCTGTCTTCT	660	
Dd	5692 CCGATACCACTTCCAGGAGAAAGAGACAGTGTGTTGACAGAAATCCACCTGTCTTCT	5751	



[illegible]

Db	5740	ACAAAGTTGGATGGAGGTGCGGTTGTCACAGGTAAGCTTCAAGGTGCAAAACCCCTTCTACGG	5799
QY	61	GAGAGGTTCAATTCCAGGTGCGGCTCAACCAATACCTGGTTGGGTGAGCTTCCATGTC	120
Db	5800	GAGAGAGGTACAAATTCTCGTTCGGGCTCAATCAATACCTGGTTGGGTGACAGCTTCCATGTC	5859
QY	121	GAGCCCGAACCAGATGTATGACAGTGTCACTTCCATGCTACCGAACCCCTCCCATATCA	180
Db	5860	GAGCCCGCAACCGGACGTAGCAGTGTCACTTCCATGCTACCGAACCCCTCCCATATACG	5919
QY	181	GCAGAGA CCGGTAAAGCGCAGGCTGGCCAGGGGGGGTCTCCCCCTCTTGGGCAAGCTTCTCA	240
Db	5920	GGGAGAACGGGTAAAGCCTTAGGCTTGCCAGAGGAGATCTCCCCCTCTTGGGCAAGCTTCTCA	5979
QY	241	GCTAGCCAGTTGTCTGAGCCCTTCTCTCGAAGCGCAGATATACCAATTAACCAAAATGACTTCCCA	300
Db	5980	GCTAGCCAGCTGTCTGAGCCCTTCTCTTGAAGCAACATGACTACCGTCAATGACTCCCGC	6039
QY	301	GACCTGAACCTCATTCGAGGCGCAACTCTCTGTGGCGGCAATGATGGGCGGGGACATTACC	360
Db	6040	GACCTGAACCTCATTCGAGGCGCAACTCTCTGTGGCGGCAATGATGGGCGGGGACATTACC	6099
QY	361	CGCGTGGAGTCCAGGAACAAGGTAGTAATCTGTGACTCTTTCGACCCGCTCCGAGCGGAG	420
Db	6100	CGCGTGGAGTCCAGGAACAAGGTAGTAATCTTGTGACTCTTTCGAGCCGCTCCGAGCGGAG	6159
QY	421	GAGAGTAGAGCGGGAAGTGTCCGTCCCGGCGGAGATCTCTGCGGAATCCAGAAATTTCCCA	480
Db	6160	GAGAGTAGAGAGGAAGTATCCGTTCCGGCGGAGATCTCTGCGGAGGTCCAGAAATTTCTCT	6219
QY	481	CCAGCGATGCCCGGATGGGCAACGCCCGGATTTACAACCTTCCGTCTGAGAGTCTTGGAGA	540
Db	6220	CGAGCGATGCCCGGATGGGCAACGCCCGGATTTACAACCTTCCGTCTGAGAGTCTTGGAGA	6279
QY	541	GGCCCGGACTACGTCCTCCAGTGGTATCATGGGAGCCCATCTGCAACCTTAAAGACCCCT	600
Db	6280	GACCCGGAATACGTCCTCCAGTGGTATCATGGGAGCCCATCTGCAACCTTAAAGACCCCT	6339
QY	601	CCTATACCACTCCACCGGAAAGAGGACAGTTGTTCTGAACAGATCCACCGTGTCTTCT	660
Db	6340	CCGATTAACAACCTCCACCGGAAAGAGGACAGTTGTTCTGATCAGAACTTACCGTGTCTTCT	6399
QY	661	GCCTTGCGGAGCTTGGCCAAAGGCTTTTGTGTGATGCTTCGGAACGTTGGCGCTTCGACAGC	720
Db	6400	GCTTTGGCGGAGCTTCGCCAAAGACCTTGGGCACTCCAAATGTTGGCGCTTCGACAGC	6459
QY	721	GGCAACGGCAACCGGCGCTCCGTAACCAATTCCTCGACGACGCGGAGGAGGATTCGAGCTT	780
Db	6460	GGCAACGGCAACCGGCGCTTCCTGACCAAGCCTTCGACGACGCGGAGGAGGATTCGAGCTT	6519
QY	781	GAGTCTGATTTCTTCATGCCCCCTTGAAGGGGAGCGGGGGGACCCCGATCTTCAGGAC	840
Db	6520	GAGTCTGATTTCTTCATGCCCCCTTGAAGGGGAGCGGGGGATTCGATCTTCAGGAC	6579
QY	841	GGGTCTTGTGTCTACCGTGAAGTGAAGAGGCGGGTGAAGAGCTGCTGTCTCTCGATGTCC	900
Db	6580	GGGTCTTGTGTCTACCGTGAAGTGAAGAGGCGGGTGAAGAGCTGCTGTCTCTCGATGTCC	6639
QY	901	TACACATGGAACGAGCGCTGATCAGCGCAGTGGCGCTGCGGAGGAAAGCAAGCTGGCCATC	960
Db	6640	TACACATGGAACGAGCGCTGATCAGCGCAGTGGCGCTGCGGAGGAAAGCAAGCTGGCCATC	6699
QY	961	AACCGTTTGAACAATCTTTTGTCTGCTGACCAACCAATGATCTACGCTACCAATCCCGC	1020
Db	6700	AATCGACTGAGCAACTTTTGTCTCGTCAACCAAACTTGGTCTATGCTACAAACATCTCGC	6759
QY	1021	AGCGCAAGCCAGCGG 1035	
Db	6760	AGCGCAAGCCAGCGG 6774	

## RESULT 7



US-10-029-907-6  
; Sequence 6, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802)...(8407)  
US-10-029-907-6

Query Match 85.5%; Score 884.6; DB 4; Length 8638;  
Best Local Similarity 90.9%; Pred. No. 3.5e-234;  
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

1 ACAGAGTGGATGGGCTGCGGCTGACAGGATGCTCGGCGTGCACAACTCTCTACGG 60  
5741 ACAGAGTGGATGGGCTGCGGCTGACAGGATGCTCGGCGTGCACAACTCTCTACGG 5800  
61 GAGGAGTGCATTCAGAGTGGGCTGACCAATGCTGGTGGGCGAGCTCCCATGC 120  
5801 GAGGAGTGCATTCAGAGTGGGCTGACCAATGCTGGTGGGCGAGCTCCCATGC 5860  
121 GAGCGCGAATCCGATGATGAGTGTCTCATCTTCATGCTCCAGACCCCTCCCATACA 180  
5861 GAGCGCGAATCCGATGATGAGTGTCTCATCTTCATGCTCCAGACCCCTCCCATAGG 5920  
181 GAGGAGAGCGCTAAGCGGAGGCTGGCCAGGGGCTTCCCTCTCTGGCCAGCTTCA 240  
5921 GCGGAGAGCGCTAAGCGGAGGCTGGCCAGGGGCTTCCCTCTCTGGCCAGCTTCA 5980  
241 GCTAGCAGTGTCTGGCCCTCTCTGAGGCGACATGATCCCAATGACTTCCCA 300  
5981 GCTAGCAGTGTCTGGCCCTCTCTGAGGCGACATGATCCCAATGACTTCCCG 6040  
301 GACGCTGACCTCATCGAGGCGCAACTCTGTGGCGGATGAGTGGGCGGACATTAC 360  
6041 GACGCTGACCTCATCGAGGCGCAACTCTGTGGCGGAGATGGGCGGAAATCACC 6100  
361 CGCGTGAAGTCAAGAAACAAGTAGTATCTGAGTCTTTGACCCGCTCGAGCGAG 420  
6101 CGCGTGAAGTCAAGAAACAAGTAGTATCTTTGAGTCTTTGACCGCTCGAAGCGAG 6160  
421 GAGGAGAGCGGGAATGTCCTCCGCGGCGGAGATCTGGGAAATTCAGAAATTTCCA 480  
6161 GAGGAGAGCGGGAATGTCCTCCGCGGCGGAGATCTGGGAGGTCAGAAATTTCCCT 6220  
481 CCAGCGATCCCGCATGAGGCGACCGCGGATTAACAACCTCCGCTGAGTCTTGAAG 540  
6221 CCAGCGATCCCGCATGAGGCGACCGCGGATTAACAACCTCCATGTTAGAGTCTTGAAG 6280  
541 GCCCGGAGTACGTCCTCTCAGTGTATCATGGTGGCCATGCTCCACTTAAGACCCCT 600  
6281 GACCCCGAGTACGTCCTCTCAGTGTATCATGGTGGCCATGCTCCACTTAAGACCCCT 6340  
601 CCTATACACCTCCAGGAGAAAGAGAGAGAGTGTCTGACAGATTCACCGGTCTTCT 660  
6341 CCGATACACCTCCAGGAGAAAGAGAGAGAGTGTCTGACAGATTCACCGGTCTTCT 6400  
661 GCGCTGGCGAGGCTTCCACAAAGGCTTTGTAGTCCGGAACCGTCCGCTGACAGC 720

6401 GCGTGGCGAGGCTCCGACAAAGACCTTCCGAGCTCCGAATCTGCGCCGTCGACAGC 6460  
721 GCGAGCGCAACCGCCCTCTCTGACCAATCTCCGAGAGCGCGAGACAGATTCGACTT 780  
6461 GCGAGCGCAACCGCCCTCTCTGACCAATCTCCGAGAGCGCGAGACAGATTCGAGT 6520  
781 GAGTGTATTCCTCCATGAGCGCCCTCTGAGGAGGAGACCGGAGATCCCATCTCAGCAG 840  
6521 GAGTGTATTCCTCCATGAGCGCCCTCTGAGGAGGAGACCGGAGATCCCATCTCAGCAG 6580  
841 GGGCTTGGTCTACCGTGAAGTGAAGAGCGCGGTGAGAGCTGTCTGCTGATGATGCC 900  
6581 GGGCTTGGTCTACCGTGAAGTGAAGAGCGCGGTGAGAGCTGTCTGCTGATGATGCC 6640  
901 TACACATGAGACAGGCGCTGTGATCAGCCATGCGCTGCGGAGAAAGAGAGTCCCATC 960  
6641 TACACATGAGACAGGCGCTGTGATCAGCCATGCGCTGCGGAGAAAGAGAGTCCCATC 6700  
961 AACGCTTGAAGCACTTTGCTGCGTCAACCAACATGATGCTACGCTACCATCCCG 1020  
6701 AATGCACTGAGCACTTTGCTGCGTCAACCAACATGATGCTACGCTACCATCCCG 6760  
1021 AGCGAAGCCAGCGG 1035  
6761 AGCGAAGCCAGCGG 6775

RESULT 8  
US-10-029-907-24  
; Sequence 24, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802)...(8407)  
US-10-029-907-24

Query Match 85.5%; Score 884.6; DB 4; Length 8638;  
Best Local Similarity 90.9%; Pred. No. 3.5e-234;  
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

1 ACAGAGTGGATGGGCTGCGGCTGACAGGATGCTCGGCGTGCACAACTCTCTACGG 60  
5741 ACAGAGTGGATGGGCTGCGGCTGACAGGATGCTCGGCGTGCACAACTCTCTACGG 5800  
61 GAGGAGTGCATTCAGAGTGGGCTGACCAATGCTGGTGGGCGAGCTCCCATGC 120  
5801 GAGGAGTGCATTCAGAGTGGGCTGACCAATGCTGGTGGGCGAGCTCCCATGC 5860  
121 GAGCGCGAATCCGATGATGAGTGTCTCATCTTCATGCTCCAGACCCCTCCCATACA 180  
5861 GAGCGCGAATCCGATGATGAGTGTCTCATCTTCATGCTCCAGACCCCTCCCATAGG 5920  
181 GAGGAGAGCGCTAAGCGGAGGCTGGCCAGGGGCTTCCCTCTCTGGCCAGCTTCA 240  
5921 GCGGAGAGCGCTAAGCGGAGGCTGGCCAGGGGCTTCCCTCTCTGGCCAGCTTCA 5980  
241 GCTAGCAGTGTCTGGCCCTCTCTGAGGCGACATGATCCCAATGACTTCCCA 300

```
Db 5981 GCTAGCAGCTGTGCGCCCTTCCTTGAAGCAATGACCTACCCGTCATCTCCCG 6040
Qy 301 GACGCTGACCTATCGAGGCAACCTCTGTGGCGGATGAGTGGCGGGGACATTACC 360
Db 6041 GACGCTGACCTATCGAGGCAACCTCTGTGGCGGATGAGTGGCGGGGACATTACC 6100
Qy 361 GCGGTGAGTCAAGAAAGAGTATGATCTGGAATCTTTTCCAGCCCGCTCCGAGGAG 420
Db 6101 CCGGTGAGTCAAGAAAGAGTATGATCTTTTCCAGCCCGCTCCGAGGAG 6160
Qy 421 GAGGATGAGCGGAAAGTGTCCGTCGGGAGATCTTCGCGAAATCCAAATTTCCCA 480
Db 6161 GAGGATGAGCGGAAAGTGTCCGTCGGGAGATCTTCGCGAAATTTCCCT 6220
Qy 481 CGAGCATGCGGATGAGGACGCGCGGATTTCAACCTTCGCTGAGAGTCTTGAAG 540
Db 6221 CGAGCATGCGGATGAGGACGCGCGGATTTCAACCTTCGCTGAGAGTCTTGAAG 6280
Qy 541 GCGCCGAGTACGTCTCCCTCAAGTGTACATGAGTGCCCATGCGCATCTAAGACCCCT 600
Db 6281 GACCGGATCTAGTCCCTCAAGTGTACATGAGTGTCCATGCGCGCTCCAAAGCCCT 6340
Qy 601 CCTATCCACTTCCAGGAAAGAGACAGTTTCTTCAAGATCTCCGCTGTTCT 660
Db 6341 CCGATACCACTTCCAGGAAAGAGACAGTTTCTTCAAGATCTCCGCTGTTCT 6400
Qy 661 GCGCTGCGGAGTTCSCCAAAAGGCTTTTGTAGTCTCCGACCGCTGCGCGTCAAGC 720
Db 6401 GCGTGTGCGGAGTTCSCCAAAAGGCTTTTGTAGTCTCCGACCGCTGCGCGTCAAGC 6460
Qy 721 GCGCAGGCAACGCGCCCTCTGATCCATCTCCGACGAGCGGAGACAGATCTGACGT 780
Db 6461 GCGCAGGCAACGCGCCCTCTGATCCATCTCCGACGAGCGGAGACAGATCTGACGT 6520
Qy 781 GAGTGTATCTTCCATGCCCCCTTGAAGGGGAGACCGCGGATCTCAAGCAGC 840
Db 6521 GAGTGTATCTTCCATGCCCCCTTGAAGGGGAGACCGCGGATCTCAAGCAGC 6580
Qy 841 GAGTGTATCTTCCATGCCCCCTTGAAGGGGAGACCGCGGATCTCAAGCAGC 900
Db 6581 GAGTGTATCTTCCATGCCCCCTTGAAGGGGAGACCGCGGATCTCAAGCAGC 6640
Qy 901 TACATGAGCAAGCGCTGTGATCAGCGCATGCGCTCGAGAGAAAGCAAGCTGCCATC 960
Db 6641 TACATGAGCAAGCGCTGTGATCAGCGCATGCGCTCGAGAGAAAGCAAGCTGCCATC 6700
Qy 961 AACGCTTGAAGCACTTTTGTGCTGCTACCAACAATGTCTACGCTACATCCGC 1020
Db 6701 AATGACTGAGCAACTTTTGTGCTGCTGCTACCAACAATGTCTACATCTCCG 6760
Qy 1021 AGCGCAAGCGAGCGG 1035
Db 6761 AGCGCAAGCGCTGCGG 6775

RESULT 9
US-10-029-907-1
; Sequence 1, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029, 907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257, 857
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 8639
; TYPE: DNA
```

```
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1803) ... (8408)
US-10-029-907-1

Query Match      85.5%; Score 884.6; DB 4; Length 8639;
Best Local Similarity 90.9%; Pred. No. 3.5e-234;
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 1 ACAGAGTGAATGAGGAGTGGGTCGCTGACAGTACCTCCGGGTGCAAACTCTCTACAGG 60
Db 5742 ACAGAGTGAATGAGGAGTGGGTCGCTGACAGTACCTCCGGGTGCAAACTCTCTACAGG 5801
Qy 61 GAGGAGTCAATTCAGAGTGGGCTCAACCAATACCTGTGTGGTGCAGCTCCATGC 120
Db 5802 GAGGAGTCAATTCAGAGTGGGCTCAACCAATACCTGTGTGGTGCAGCTCCATGC 5861
Qy 121 GAGCCGGAACCGAGTGTAGAGTGTCACTTTCATGTCTACCGACCCCTCCCAATCA 180
Db 5862 GAGCCGGAACCGAGTGTAGAGTGTCACTTTCATGTCTACCGACCCCTCCCAATCA 5921
Qy 181 GCAAGAGAGCTTAAGCGAGGCTGGCCAGGGGATCTCCCGCTCTTGGCCAGCTTTCA 240
Db 5922 GCGAGAGAGCTTAAGCGAGGCTGGCCAGGGGATCTCCCGCTCTTGGCCAGCTTTCA 5981
Qy 241 GCTAGCCAGTGTCTGAGCCCTTCTCGAAGGCGACATACATTACCAAAATGACTTCCCA 300
Db 5982 GCTAGCCAGTGTCTGAGCCCTTCTCGAAGGCGACATACATTACCAAAATGACTTCCCA 6041
Qy 301 GACGCTGACCTCATGAGAGGCAACCTCTGTGCGGAGTGAAGATGGCGGGGACATTAC 360
Db 6042 GACGCTGACCTCATGAGAGGCAACCTCTGTGCGGAGTGAAGATGGCGGGGACATTAC 6101
Qy 361 GCGGTGAGTCAAGAGAAAGATAGTAACTCTGAGACTTTTTCGACCCGCTCCGAGGAG 420
Db 6102 GCGGTGAGTCAAGAGAAAGATAGTAACTTTTTCGAGCCGCTCCGAGGAG 6161
Qy 421 GAGGATGAGCGGAAAGTGTCCGTCGCGCGGAGATCTCTGCGGAAATCCAAAGAAATTTCCCA 480
Db 6162 GAGGATGAGCGGAAAGTGTCCGTCGCGCGGAGATCTCTGCGGAAATTTCCCT 6221
Qy 481 CGAGGATGCGCGAGTGGGACGCGCGGATTTCAACCTCTGCTGCTGAGTCTTGAAG 540
Db 6222 CGAGGATGCGCGAGTGGGACGCGCGGATTTCAACCTCTGCTGAGTCTTGAAG 6281
Qy 541 GCGCCGAGTCAAGTCCCTTCCAGTGTATAGGAGTCCCACTGCGCATCTAAGACCCCT 600
Db 6282 GACCGGAGTCAAGTCCCTTCCAGTGTATAGGAGTCCCACTGCGCATCTAAGACCCCT 6341
Qy 601 CCTATACCACTTCCAGGAAAGAGACAGTGTGTGACAGAAATCCACGTCCTTCT 660
Db 6342 CCGATACCACTTCCAGGAAAGAGACAGTGTGTGACAGAAATCCACGTCCTTCT 6401
Qy 661 GCGCTGCGGAGCTTTCGCAAAAGCTTTTGTGATCTCGAACGTCGCGCTGACAGC 720
Db 6402 GCGTGTGCGGAGCTTTCGCAAAAGCTTTCGAGACTCCGAATGTCGCGCTGACAGC 6461
Qy 721 GGCACGAGCAACCGCCCTCTCTGACCAATCTCTCCAGACGCGCGAGAGATCTGACGTT 780
Db 6462 GGCACGAGCAACCGCCCTCTCTGACCAATCTCTCCAGACGCGCGAGATCTGACGTT 6521
Qy 781 GAGTGTATCTTCCATGCCCCCTTGAAGGGGAGACCGGGGATCCCGATCTGAGCAGC 840
Db 6522 GAGTGTATCTTCCATGCCCCCTTGAAGGGGAGACCGGGGATCCCGATCTGAGCAGC 6581
Qy 841 GCGTGTGCTTACCGTGAAGTGAAGAGCGCGTGAAGACGTCGTCGTCGATGTCC 900
Db 6582 GCGTGTGCTTACCGTGAAGTGAAGAGCGCGTGAAGACGTCGTCGTCGATGTCC 6641
Qy 901 TACATGAGCAAGCGCTCTGATCAAGCATGCGCTGCGAGAGAAAGCAAGTGTCCATC 960
Db 6642 TACATGAGCAAGCGCTCTGATCAAGCATGCGCTGCGAGAGAAAGCAAGTGTCCATC 6701
```



APPLICANT: Bartschlagel, Ralf FW  
FILE REFERENCE: all sequences  
CURRENT APPLICATION NUMBER: US/09/539,601C  
CURRENT FILING DATE: 2001-08-30  
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 11076  
TYPE: DNA  
ORGANISM: Hepatitis C virus  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: (1)..(341)  
OTHER INFORMATION: construct I389/Core-3'/wt  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (342)..(1193)  
OTHER INFORMATION: hepatitis C virus core - neomycin  
OTHER INFORMATION: phosphotransferase fusion protein  
FEATURE:  
NAME/KEY: RBS  
LOCATION: (1202)..(1812)  
OTHER INFORMATION: internal ribosome entry site from  
OTHER INFORMATION: encephalomyocarditis virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1813)..(10845)  
OTHER INFORMATION: hepatitis C virus polypeptide from core to  
OTHER INFORMATION: nonstructural protein NS5B; parental sequence  
OTHER INFORMATION: without cell culture-adaptive mutations  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: (10846)..(11076)  
US-09-539-601-1

Query Match 85.5%; Score 884.6; DB 4; Length 11076;  
Best Local Similarity 90.9%; Pred. No. 3.8e-234;  
Matches 941; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1 ACAGAGTGAATGGGGTGGCTGACAGGTAGCTCCGGCGTGCAACCTCTCTACCG 60  
DB 8179 ACAGAGTGAATGGGGTGGCTGACAGGTAGCTCCGGCGTGCAACCTCTCTACCG 8238  
QY 61 GAGGAGTCACTTCAGGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCATCC 120  
DB 8239 GAGGAGTCACTTCAGGTGGGCTCAACCAATACCTGGTGGTGGAGCTCCATCC 8298  
QY 121 GAGCCCGAAGCGGATGATGAGTCACTTCAGGTGGTGGTGGTGGTGGTGGTGG 180  
DB 8299 GAGCCCGAAGCGGATGATGAGTCACTTCAGGTGGTGGTGGTGGTGGTGGTGG 8358  
QY 181 GCAGAGACGGCTAAGCGAGGTGGCGAAGGGGTCTCCCTCTCTGGCCAGCTCTTCA 240  
DB 8359 GCAGAGACGGCTAAGCGAGGTGGCGAAGGGGTCTCCCTCTCTGGCCAGCTCTTCA 8418  
QY 241 GCTAGCCAGTGTCTGGCCCTTCTCTGAGAGCGACATTAATCAATTAATCACTTCCA 300  
DB 8419 GCTAGCCAGTGTCTGGCCCTTCTCTGAGAGCGACATTAATCAATTAATCACTTCCA 8478  
QY 301 GACGCTGACCTCATCGAGGCGAAGCTCTGAGCGGCGATGAGAGTGGGGGAGCATTAAC 360  
DB 8479 GACGCTGACCTCATCGAGGCGAAGCTCTGAGCGGCGATGAGAGTGGGGGAGCATTAAC 8538  
QY 361 CGCGTGAAGTCAAGAAACAAGTATATCTGAGCTCTTCTGACCCCGCTCGAGCGAG 420  
DB 8539 CGCGTGAAGTCAAGAAACAAGTATATCTGAGCTCTTCTGAGCGCGCTCGAAGCGAG 8598  
QY 421 GAGGATGACGGGAAGTGTCCGCTCGCGGAGAGTCTCGCGAATTCAGAAATTCCTCA 480  
DB 8599 GAGGATGACGGGAAGTGTCCGCTCGCGGAGAGTCTCGCGAATTCAGAAATTCCTCA 8658

QY 481 CCAGGATGCCCGATGGGACGCGCCGGATTACAACCTCCGCTGTGAGTCTGGAAG 540  
DB 8659 CCAGGATGCCCGATGGGACGCGCCGGATTACAACCTCCGCTGTGAGTCTGGAAG 8718  
QY 541 GCCCGGACCTAGTCCCTCCAGTGTATAGTGGTCCCACTCCCACTTAAGACCTT 600  
DB 8719 GCCCGGACCTAGTCCCTCCAGTGTATAGTGGTCCCACTCCCACTTAAGACCTT 8778  
QY 601 CCTATACCACTCCACCGAAGAAAGGACATTTGTTGAGAGATTCACCGTCTTCT 660  
DB 8779 CCTATACCACTCCACCGAAGAAAGGACATTTGTTGAGAGATTCACCGTCTTCT 8838  
QY 661 GCCCTGGGAGACTTGGCAAAAGCTTTGTAGTCTCCGACCGTCCGCGGACGAC 720  
DB 8839 GCCCTGGGAGACTTGGCAAAAGCTTTGTAGTCTCCGACCGTCCGCGGACGAC 8898  
QY 721 GGCACGGCAACCGCCCTCTGACCAATCTTCCGACGAGCGGACGAGATCTGACGTT 780  
DB 8899 GGCACGGCAACCGCCCTCTGACCAATCTTCCGACGAGCGGACGAGATCTGACGTT 8958  
QY 781 GAGTGTATTTCTCTCATGCCCCCTTGAAGGGAGGACCGGGGACCCCGATCTGACGAC 840  
DB 8959 GAGTGTATTTCTCTCATGCCCCCTTGAAGGGAGGACCGGGGACCCCGATCTGACGAC 9018  
QY 841 GGGTCTGTATACCGTGAAGGAGGCGGAGGAGCTGTCTGTCTGCTGATGTCC 900  
DB 9019 GGGTCTGTATACCGTGAAGGAGGCGGAGGAGCTGTCTGTCTGCTGATGTCC 9078  
QY 901 TACACATGACAGGCGCTGTATACGCGCATGCGTGGAGAGAAAGCAAGCTGCCATC 960  
DB 9079 TACACATGACAGGCGCTGTATACGCGCATGCGTGGAGAGAAAGCAAGCTGCCATC 9138  
QY 961 AACGCTTGAACCACTTTTCTGTGTGATCAACAATGTCTACGATCAATCCCGC 1020  
DB 9139 AACGCTTGAACCACTTTTCTGTGTGATCAACAATGTCTACGATCAATCCCGC 9198  
QY 1021 AGCGCAAGCCAGCGG 1035  
DB 9199 AGCGCAAGCCAGCGG 9213

RESULT 12  
US-10-029-907-7  
Sequence 7, Application US/10029907  
Patent No. 6706874  
GENERAL INFORMATION:  
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
FILE REFERENCE: 13/083  
CURRENT APPLICATION NUMBER: US/10/029,907  
PRIOR FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/257,857  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 8638  
TYPE: DNA  
ORGANISM: HCV  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1802)...(8407)  
US-10-029-907-7

Query Match 85.3%; Score 883; DB 4; Length 8638;  
Best Local Similarity 90.8%; Pred. No. 9.7e-234;  
Matches 940; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1 ACAGAGTGAATGGGGTGGCTGACAGGTAGCTCCGGGCTGCAAACTCTCTACCG 60  
DB 5741 ACAGAGTGAATGGGGTGGCTGACAGGTAGCTCCAGGCTGCAAACTCTCTACCG 5800

61 GAGAGAGTCAATTCAGAGTGGGGCTCAACCAATACCTGGTGGTGGAGCTCCATG 120  
DB 5801 GAGAGAGTCAATTCAGAGTGGGGCTCAACCAATACCTGGTGGTGGAGCTCCATG 5860  
QY 121 GAGCCGGAACCGGATGAGAGTGTCACTTCCATGCTCAACGACCCCTCCCATGACA 180  
DB 5861 GAGCCGGAACCGGATGAGAGTGTCACTTCCATGCTCAACGACCCCTCCCATGACA 5920  
QY 181 GAGAGAGTCAATTCAGAGTGGGGCTCAACCAATACCTGGTGGTGGAGCTCCATG 240  
DB 5921 GAGAGAGTCAATTCAGAGTGGGGCTCAACCAATACCTGGTGGTGGAGCTCCATG 5980  
QY 241 GCTAGGCACTGTGTGGGCTTCTTCCGAGGAGCAATACCTTCAACCAATGACTTCCA 300  
DB 5981 GCTAGGCACTGTGTGGGCTTCTTCCGAGGAGCAATACCTTCAACCAATGACTTCCA 6040  
QY 301 GAGCTGAGCTTCAATTCAGAGGCAACCTCTGTGGCGGATGAGATGGGCGGGAATAC 360  
DB 6041 GAGCTGAGCTTCAATTCAGAGGCAACCTCTGTGGCGGATGAGATGGGCGGGAATAC 6100  
QY 361 CGGCTGAGTCAAGAGCAAGATGATGATCTGTGACTCTTTCGACCCGCTCCGAGCGAG 420  
DB 6101 CGGCTGAGTCAAGAGCAAGATGATGATCTGTGACTCTTTCGACCCGCTCCGAGCGAG 6160  
QY 421 GAGGATGAGCGGGAATGTCCCTCCGCGGAGATCTGTGGGAAATTCAGAAATTCCTCA 480  
DB 6161 GAGGATGAGCGGGAATGTCCCTCCGCGGAGATCTGTGGGAAATTCAGAAATTCCTCA 6220  
QY 481 CCAGCATGATCCGATGAGGAGCGCGGATTAACCTCTCGCTGCTGAGTCTCTGAG 540  
DB 6221 CCAGCATGATCCGATGAGGAGCGCGGATTAACCTCTCGCTGCTGAGTCTCTGAG 6280  
QY 541 GCCCGGAGTACGCTCCCTCCAGTGTGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 6281 GACCCGAGTACGCTCCCTCCAGTGTGATGATGATGATGATGATGATGATGATGATGAT 6340  
QY 601 CCTATTCACCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 6341 CGGATTCACCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6400  
QY 661 GCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
DB 6401 GCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6460  
QY 721 GGCACGGAACCGGCTCTCTGACCAATCTCTGACGAGCGGCGAGGAGGATCTGACGTT 780  
DB 6461 GGCACGGAACCGGCTCTCTGACCAATCTCTGACGAGCGGCGAGGAGGATCTGACGTT 6520  
QY 781 GAGTGTATTTCTTCCATGCCCCCTTGAAGGGGAGCGGGGAGCCCGATCTCAGCGAC 840  
DB 6521 GAGTGTATTTCTTCCATGCCCCCTTGAAGGGGAGCGGGGAGCCCGATCTCAGCGAC 6580  
QY 841 GGGCTTGTGCTTACCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
DB 6581 GGGCTTGTGCTTACCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6640  
QY 901 TACACATGAGCAGGCGCTGTGATCAGCGCATGCGCTGCGAGGAGGAGGAGGAGGAGGAG 960  
DB 6641 TACACATGAGCAGGCGCTGTGATCAGCGCATGCGCTGCGAGGAGGAGGAGGAGGAGGAG 6700  
QY 961 AACGCTTGAAGCACTTTTGTGCTGCTCAGCAGCAATGATCTTACGCTACATCCCG 1020  
DB 6701 AACGCTTGAAGCACTTTTGTGCTGCTCAGCAGCAATGATCTTACGCTACATCCCG 6760  
QY 1021 AGCGCAAGCGAGCGG 1035  
DB 6761 AGCGCAAGCGAGCGG 6775

RESULT 13  
US-10-029-907-25  
; Sequence 25, Application us/10029907

Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: HEPATITIS C VIRUS  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802)...(8407)  
US-10-029-907-25

Query Match 85.3%; Score 883; DB 4; Length 8638;  
Best Local Similarity 90.8%; Pred. No. 9,7e-234;  
Matches 940; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1 ACAGAGTGAATGGGATGGGCTGACAGATAGCTCCGCGCTGCAAACTCTCTTACGG 60  
DB 5741 ACAGAGTGAATGGGATGGGCTGACAGATAGCTCCGCGCTGCAAACTCTCTTACGG 5800  
QY 61 GAGAGAGTCAATTCAGAGTGGGGCTCAACCAATACCTGTGGTGGAGCTCCATG 120  
DB 5801 GAGAGAGTCAATTCAGAGTGGGGCTGACAGATAGCTCCGCGCTGCAAACTCTCTTACGG 5860  
QY 121 GAGCCGGAACCGGATGAGAGTGTCTCACTTCATGCTCAACGACCCCTCCCATGACA 180  
DB 5861 GAGCCGGAACCGGATGAGAGTGTCTCACTTCATGCTCAACGACCCCTCCCATGACA 5920  
QY 181 GAGAGAGTCAATTCAGAGTGGGGCTCAACCAATACCTGTGGTGGAGCTCCATG 240  
DB 5921 GAGAGAGTCAATTCAGAGTGGGGCTGACAGATAGCTCCGCGCTGCAAACTCTCTTACGG 6160  
QY 241 GCTAGGCACTGTGTGGGCTTCTTCCGAGGAGCAATACCTTCAACCAATGACTTCCA 300  
DB 5981 GCTAGGCACTGTGTGGGCTTCTTCCGAGGAGCAATACCTTCAACCAATGACTTCCA 6040  
QY 301 GAGCTGAGCTTCAATTCAGAGGCAACCTCTGTGGCGGATGAGATGGGCGGGAATAC 360  
DB 6041 GAGCTGAGCTTCAATTCAGAGGCAACCTCTGTGGCGGATGAGATGGGCGGGAATAC 6100  
QY 361 CGGCTGAGTCAAGAGCAAGATGATGATCTGTGACTCTTTCGACCCGCTCCGAGCGAG 420  
DB 6101 CGGCTGAGTCAAGAGCAAGATGATGATCTGTGACTCTTTCGACCCGCTCCGAGCGAG 6160  
QY 421 GAGGATGAGCGGGAATGTCCCTCCGCGGAGATCTGTGGGAAATTCAGAAATTCCTCA 480  
DB 6161 GAGGATGAGCGGGAATGTCCCTCCGCGGAGATCTGTGGGAAATTCAGAAATTCCTCA 6220  
QY 481 CCAGCATGATCCGATGAGGAGCGCGGATTAACCTCTCGCTGCTGAGTCTCTGAG 540  
DB 6221 CCAGCATGATCCGATGAGGAGCGCGGATTAACCTCTCGCTGCTGAGTCTCTGAG 6280  
QY 541 GCCCGGAGTACGCTCCCTCCAGTGTGATCAGTGGTCCCATGCGACTTAAAGCCCT 600  
DB 6281 GAGCCGAGTACGCTCCCTCCAGTGTGATCAGTGGTGTGCTCAAGATCTTCCATCC 6340  
QY 601 CCTATTCACCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 6341 CGGATTCACCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6400  
QY 661 GCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
DB 6401 GCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6460

QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGACGGGAGACGATCTGACGTT 780  
DB 6461 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGACGGGAGACGATCTGACGTT 6520  
QY 781 GAGTCGATTTCTCCATGCCCCCTTGAAGGGAGACCGGGAGACCCGATCTCAGCGAC 840  
DB 6521 GAGTCGATTTCTCCATGCCCCCTTGAAGGGAGACCGGGAGATCTCCGATCTCAGCGAC 6580  
QY 841 GGGTCGTTGCTACCGGTGATGAGGAGCGCGGTGAGGAGCGTCTGCTGATGATGTC 900  
DB 6581 GGGTCGTTGCTACCGGTGATGAGGAGCGGTGAGGAGCGTCTGCTGATGATGTC 6640  
QY 901 TACACATGACAGCGCCCTCTGATCAAGCCATGCGCTCGGAGAGAAAGCAAGCTGCCATC 960  
DB 6641 TACACATGACAGCGCCCTCTGATCAAGCCATGCGCTCGGAGAGAAAGCAAGCTGCCATC 6700  
QY 961 AACGCGTGAAGCACTTTTGTGCTGACACACACATGATGCTACGCTACCATCCGC 1020  
DB 6701 AATGCACTGAGCAACTTTTGTGCTGCTGACACCAACTGCTATGCTACCAATCTCCG 6760  
QY 1021 AGCGCAAGCGACGGG 1035  
DB 6761 AGCGCAAGCGCTGGG 6775

RESULT 14  
US-10-029-907-4  
Sequence 4, Application US/10029907  
Patent No. 6706874  
GENERAL INFORMATION:  
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
HEPATITIS C VIRUS  
FILE REFERENCE: 13/083  
CURRENT APPLICATION NUMBER: US/10/029,907  
PRIOR FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/257,857  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 8643  
TYPE: DNA  
ORGANISM: HCV  
FEATURE:  
NAME/KEY: CBS  
LOCATION: (1802) ... (8407)  
US-10-029-907-4

Query Match 85.3%; Score 883; DB 4; Length 8643;  
Best Local Similarity 90.8%; Pred. No. 9,7e-234;  
Matches 940; Conservative 0; Mismatches 95; Indels 0; Gaps 0;  
QY 1 ACAGAAAGTGAAGGGGTGCGGCTGCAAGATGAGCTCCGCGGTGCAAACTCTCTTACGG 60  
DB 5741 ACAGAAAGTGAAGGGGTGCGGCTGCAAGATGAGCTCCGCGGTGCAAACTCTCTTACGG 5800  
QY 61 GAGGAGTCACTTCCAGGTGCGGCTCAACCAATACCTGTTGGTTCGAGCTCCCATC 120  
DB 5801 GAGGAGTCACTTCCAGGTGCGGCTCAACCAATACCTGTTGGTTCGAGCTCCCATC 5860  
QY 121 GAGCCGAACCGGATGATGAGGAGTCACTTCCATGCTCCAGACCCCTCCCATCA 180  
DB 5861 GAGCCGAACCGGATGATGAGGAGTCACTTCCATGCTCCAGACCCCTCCCATCA 5920  
QY 181 GCAGAGACGGCTAAGCGCAGGCTGCGCAGGGGGTCTCCCTCTCTTGGCAGCTCTTCA 240  
DB 5921 GCAGAGACGGCTAAGCGCAGGCTGCGCAGGGGGTCTCCCTCTCTTGGCAGCTCTTCA 5980  
QY 241 GCTAGCCAGTTGCTGCGCTTCTCTGAAAGCGACATTAACCAAAATGACTTCCA 300  
DB 5981 GCTAGCCAGTTGCTGCGCTTCTCTGAAAGCGACATTAACCAAAATGACTTCCA 6040

QY 301 GACGCTGACCTCATGAGGCGCAACCTCTGTTGGCGGACATGAGATGGCGGGACATTACC 360  
DB 6041 GACGCTGACCTCATGAGGCGCAACCTCTGTTGGCGGACATGAGATGGCGGGACATTACC 6100  
QY 361 GCGGTGAGTCAAGAGAACAGATAGTAATCTGGAATCTTTCCAGCCGCTCCGAGCGAG 420  
DB 6101 GCGGTGAGTCAAGAGAACAGATAGTAATCTGGAATCTTTCCAGCCGCTCCGAGCGAG 6160  
QY 421 GAGGATGACCGGAGAGTGTCCGTCGCGCGGAGATCTCTGCGGAATCCAAATAATCCCA 480  
DB 6161 GAGGATGACCGGAGAGTGTCCGTCGCGCGGAGATCTCTGCGGAATCCAAATAATCCCA 6220  
QY 481 CGAGGATGCCCGGATGAGGAGCGCCGAGATTACACCTCTCGCTGAGTCTCTGAG 540  
DB 6221 CGAGGATGCCCGGATGAGGAGCGCCGAGATTACACCTCTCGCTGAGTCTCTGAG 6280  
QY 541 GCGCGGATCACTGCTCTCTGAGTATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 6281 GACCGGATCACTGCTCTCTGAGTATGATGATGATGATGATGATGATGATGATGATGAT 6340  
QY 601 CCTATACCACTCCACGAGAGAGAGACAGTGTCTGACAGATCCACCTGCTCTCT 660  
DB 6341 CGATACCACTCCACGAGAGAGAGACAGTGTCTGACAGATCCACCTGCTCTCT 6400  
QY 661 GCGCTGCGGAGCTTGCACAAAGGCTTTGATGCTCCGAGCCGTCGCGAGCAGC 720  
DB 6401 GCGCTGCGGAGCTTGCACAAAGGCTTTGATGCTCCGAGCCGTCGCGAGCAGC 6460  
QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCCAGACGCGGAGAGATCTGACGTT 780  
DB 6461 GGCACGGCAACCGCCCTCTCTGACCAATCTCTCCAGACGCGGAGAGATCTGACGTT 6520  
QY 781 GAGTCGATTTCTCCATGCCCCCTTGAAGGGAGACCGGGAGATCTCCGATCTCAGCGAC 840  
DB 6521 GAGTCGATTTCTCCATGCCCCCTTGAAGGGAGACCGGGAGATCTCCGATCTCAGCGAC 6580  
QY 841 GGGTCGTTGCTACCGGTGATGAGGAGCGCGGTGAGGAGCGTCTGCTGATGATGTC 900  
DB 6581 GGGTCGTTGCTACCGGTGATGAGGAGCGCGGTGAGGAGCGTCTGCTGATGATGTC 6640  
QY 901 TACACATGACAGCGCCCTCTGATCAAGCCATGCGCTCGGAGAGAAAGCAAGCTGCCATC 960  
DB 6641 TACACATGACAGCGCCCTCTGATCAAGCCATGCGCTCGGAGAGAAAGCAAGCTGCCATC 6700  
QY 961 AACGCGTGAAGCACTTTTGTGCTGACACACATGATGCTACGCTACCATCCGC 1020  
DB 6701 AATGCACTGAGCAACTTTTGTGCTGCTGACACCAACTGCTATGCTACCAATCTCCG 6760  
QY 1021 AGCGCAAGCGACGGG 1035  
DB 6761 AGCGCAAGCGCTGGG 6775

RESULT 15  
US-10-029-907-5  
Sequence 5, Application US/10029907  
Patent No. 6706874  
GENERAL INFORMATION:  
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
HEPATITIS C VIRUS  
FILE REFERENCE: 13/083  
CURRENT APPLICATION NUMBER: US/10/029,907  
PRIOR FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/257,857  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 8648  
TYPE: DNA  
ORGANISM: HCV  
FEATURE:

NAME/KEY: CDS  
LOCATION: (1802)... (8407)  
US-10-023-907-5

Query Match 85.3%; Score 883; DB 4; Length 8648;  
Best Local Similarity 90.8%; Pred. No. 9, 7e-234;  
Matches 940; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
QY 1 AAGAGAGTGGAGTGGGCTGGCTGCAAGGTAGCGTCCGGCGTGCAGAACTCTCTCCACGG 60
DB 5741 AAGAGAGTGGAGTGGGCTGGCTGCAAGGTAGCGTCCGGCGTGCAGAACTCTCTCCACGG 5800
QY 61 GAGAGAGTCACTTCAGGTCGGGCTCAACCAATACCTGGTGGTGGAGCTCCATGC 120
DB 5801 GAGAGAGTCACTTCAGGTCGGGCTCAACCAATACCTGGTGGTGGAGCTCCATGC 5860
QY 121 GAGCCCGAACCAGATAGTACAGTCTCACTTCATGCTGACCGACCCCTCCACATCACA 180
DB 5861 GAGCCCGAACCAGATAGTACAGTCTCACTTCATGCTGACCGACCCCTCCACATCACA 5920
QY 181 GAGAGAGCGGCTAAGCGAGGCTGGCGAGGGGGTCTCCCGCTCTTGGCGAGCTTCA 240
DB 5921 GCGAGACCGCTAAGCGAGGCTGGCGAGGGGGTCTCCCGCTCTTGGCGAGCTTCA 5980
QY 241 GCTAGCCAGTGTCTGGCCCTTCTCGAAGGCGACATATACCCAAATGACTTCCCA 300
DB 5981 GCTAGCCAGTGTCTGGCCCTTCTCGAAGGCGACATATACCCCTCAATGACTTCCCG 6040
QY 301 GACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATAGATGGGCGGGGACATTAAC 360
DB 6041 GAGGCTGACCTCATCGAGGCCAACCTCTGTGGCGGAGATAGATGGGCGGGACATCAC 6100
QY 361 CGCGTGGAGTCAAGAGAACAGGTAGTATCTGCACTCTTTCGACCCGCTCCAGCGAG 420
DB 6101 CGCGTGGAGTCAAGAGAACAGGTAGTATCTTGGACTCTTTCGAGCCGCTCCAGCGAG 6160
QY 421 GAGGATGAGCGGGAAGTGTCCGTCGGGGAGATCTCGGGAATCCAGAAATCCCA 480
DB 6161 GAGGATGAGCGGGAAGTGTCCGTCGGGGAGATCTCGGGAATCCGAGAAATCCCT 6220
QY 481 CCAGCGATGCCGATGAGGACGCGCGGATTAACACCTCCGCTGCTGAGTCTGAAAG 540
DB 6221 CGAGCGATGCCGATGAGGACGCGCGGATTAACACCTCCGCTGAGTCTGAAAG 6280
QY 541 GCCCGGAGTACGTCTCCCTCAGTGTACATGGGTGCCACTGCGACCTTACTAAGACCCCT 600
DB 6281 GACCCCGAGTACGTCTCCCTCAGTGTACATGGGTGTCCATTTGCGCTGCGCAAGCCCT 6340
QY 601 CCTATACCACTCCAGGAGAAAGAGACAGTTGTTCTGACAGATCCACCGTCTTCT 660
DB 6341 CCGATACCACTCCAGGAGAAAGAGACAGTTGTTCTGACAGATCTACCGTCTTCT 6400
QY 661 GCCCTGGCGAGGCTTCCACAAAGCTTTTGTAGCTCCGACCGTCCGCGCTGACAGC 720
DB 6401 GCTTGGCGAGGCTTCCACAAAGCTTCCGACGCTCCGAATCTGCGCTGACAGC 6460
QY 721 GGCACGGCAACCGCCCTCTCTGACCAATCTCCGACGCGGAGCAGAGATCTGACGTT 780
DB 6461 GGCACGGCAACCGCCCTCTCTGACCAAGCCTCCGACGCGGAGCAGAGATCTGACGTT 6520
QY 781 GAGTGTATTTCTCCATGCCCCCTTGAAGGGGAGACCGGGGATCCCGATCTCAGCGAC 840
DB 6521 GAGTGTATTTCTCCATGCCCCCTTGAAGGGGAGACCGGGGATCCCGATCTCAGCGAC 6580
QY 841 GGGTCTTGGTCTACCGTGAAGGAGCGCGGTGAGAGCTGCTGCTGCTGATGTC 900
DB 6581 GGGTCTTGGTCTACCGTGAAGGAGCGCGGTGAGAGCTGCTGCTGCTGATGTC 6640
QY 901 TACACATGAGCAGGCGCTGTGATCACGCCATGCGCTGCGAGGAAAGCAAGCTGCCATC 960
DB 6641 TACACATGAGCAGGCGCTGTGATCACGCCATGCGCTGCGAGGAAAGCAAGCTGCCATC 6700
QY 961 AACGCTTGAACACTTTGCTGCGTCAACAACAATGATGTACGCTACCAATCCCGC 1020
```

```
DB 6701 AATGACTAGCAACTTTGCTCCGTACCACTTGCTATGCTACACATCTCCG 6760
QY 1021 AGCGCAGCGCAGCGG 1035
DB 6761 AGCGCAGCGCAGCGG 6775
```

Search completed: February 19, 2005, 11:51:38  
Job time : 181.108 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2005, 00:03:16 ; Search time 3612.18 Seconds  
(without alignments)  
10906.573 Million cell updates/sec

Title: US-09-664-363-4

Perfect score: 1035  
Sequence: 1 ACAGAGTGTGATGGGTGGG.....CCCCAGCGCAGCCAGCGG 1035

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 14239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	49.2	4.8	925	9	CNS0091P
C 2	45.4	4.4	914	2	BF257968 HVSMEF001
C 3	44.4	4.3	486	4	BG558244 RH122_65
C 4	44.4	4.3	1101	9	CNS0175Y
C 5	44.4	4.3	2100	9	AL108460 Drosophila
C 6	44.4	4.3	506	5	CL945925 OaIFSB001
C 7	44.4	4.3	506	7	BP153627
C 8	43.4	4.2	587	6	CN010969
C 9	43.4	4.2	487	7	CA679773 WHB3878_D
C 10	43.4	4.2	540	6	CV039737 wim4.pk00
C 11	43.4	4.2	573	6	CD727100 4031393_1
C 12	43.4	4.2	591	4	CD216293 ppp2n.pk0
C 13	43.4	4.2	601	4	BI391789
C 14	43.4	4.2	615	4	BM425602 p9f2c.pk0
C 15	43.4	4.2	615	4	BM425623 p9f2c.pk0
C 16	43.4	4.2	665	1	AJ455557
C 17	43.4	4.2	667	1	AJ455557
C 18	43.4	4.2	673	7	AJ450427
C 19	43.4	4.2	674	1	CO507218
C 20	43.4	4.2	690	1	AJ449106
C 21	43.4	4.2	695	4	AJ445947
C 22	43.4	4.2	698	5	BM486338
C 23	43.4	4.2	704	6	BU481178
C 24	43.4	4.2	727	6	CD763744

25	43	4.2	737	5	BU281256
26	43	4.2	771	7	CN225007
27	43	4.2	789	7	CN217189
28	43	4.2	830	7	CN232609
29	43	4.2	866	1	AJ447521
30	42.4	4.1	1005	6	CA184993
31	42	4.1	595	1	AI980698
32	41.8	4.0	573	1	AA518979
33	41.8	4.0	584	1	AV922850
34	41.4	4.0	580	1	AV922834
35	41	4.0	673	1	AJ452584
36	41	4.0	679	7	CN140848
37	40.8	3.9	628	9	CL953354
38	40.8	3.9	628	6	CA152295
39	40.8	3.9	828	6	CA141234
40	40.6	3.9	514	7	CN433113
41	40.4	3.9	835	2	BF256551
42	40.2	3.9	420	1	AJ463083
43	40	3.9	509	4	BI996634
44	40	3.9	595	5	BO472424
45	40	3.9	634	5	BO472919

#### ALIGNMENTS

RESULT 1  
CNS0091P/c  
LOCUS  
DEFINITION  
CNS0091P 925 bp DNA linear GSS 03-JUN-1999  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION  
AL053013.1 GI:4934461  
VERSION  
AL053013.1  
KEYWORDS  
GSS.  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
Direct Submission  
Submitted 1 (bases 1 to 925)  
Submitted 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
Web : www.genoscope.cns.fr)

COMMENT  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

#### FEATURES

source  
1..925  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR19D16"  
/clone\_lib="RPCI-98"  
/note="end : TET3"

#### ORIGIN

Query Match 4.8%; Score 49.2; DB 9; Length 925;  
Best Local Similarity 15.6%; Pred. No. 0.082;



Fax: 706 583 0210  
Email: mmp@prc.uga.edu  
Seq primer: JEM REV  
High quality sequence stop: 345  
POLYA=No.

## FEATURES

source

Location/Qualifiers  
1. 486  
/organism="Sorghum propinquum"  
/mol\_type="rRNA"  
/db\_xref="taxon:132711"  
/note="Organ: Rhizomes; Vector: pBlueScript II from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

## ORIGIN

Query Match 4.3%; Score 44.4; DB 4; Length 486;  
Best Local Similarity 54.2%; Pred. No. 1.4;  
Matches 90; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
QY 720 CCGACGCGCAACCGCCCTCTTCAATCTTCGACGACGCGGAGAGATCTGACGT 779  
DB 22 CCGTACCGCCACCGCATGCGCCGACCATCCGCGCGCGCGCTCAACCGCATGACGT 81  
QY 780 TGAAGTGTATTCCTCCATCGCCCGCTTGAAGGGGAGCGCGGGAGCCCGATCTCAGCA 839  
DB 82 CCGGCGCGCTGCGCTCTTCCACCGCTGCGCGGAGAGCGCGCGCGCGCATGGA 141  
QY 840 CCGGCTCTTGTATCCGTCGAGAGAGCGCGGTGAGCACTGCTGTC 885  
DB 142 GGAAGCGCGGCGCGCGCGGAGAGGTGCGCGGTGCGCGCGGAC 187

RESULT 4  
LOCUS CNS0175Y 1101 bp DNA linear GSS 26-JUL-1999  
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC  
BACN37JL08 of DrosBAC library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION AL108460.1 GI:5628764  
VERSION AL108460  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.

REFERENCE Direct Submission  
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billard at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

## FEATURES

source

Location/Qualifiers  
1. 1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone\_lib="BACN37JL08"  
/clone\_lib="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : Sp6"

## ORIGIN

Query Match 4.3%; Score 44.4; DB 9; Length 1101;  
Best Local Similarity 14.8%; Pred. No. 1.5;  
Matches 61; Conservative 176; Mismatches 175; Indels 0; Gaps 0;  
QY 61 GAGAGTCACATTCAGGTCGGGCTCAACATATCTGTGGTGGCAGCTCCATGC 120  
DB 675 GKKAKGVGRGIVCCAGGAGACACMAADCGCCACMACCCSSSSSCASTSSAS 734  
QY 121 GAGCCCGAACCAGATTATACATGCTCATCTTCATCTCAACGACCCCTCCACATCAC 180  
DB 735 RGVVSSCACSSGSGGAGACAGAGGSGGGGCGCCAGSSGVCAACGACSSASMC 794  
QY 181 GCAGAGACCGCTTAAGGCGAGGCTGCGCAGGGGTCTCCCTCTTGGCCAGCTTTCA 240  
DB 795 SVSSGSSCCASCCSCCTGVSSCAVBSASASVMSKVASAVACSAVASGMSAGVSSGR 854  
QY 241 GCTAGCCAGTTGTCTGGCCCTTCTCGAAGCGCATACATTACCCAAATGACTTCCA 300  
DB 855 SSVASSVSAASVSSSSSSSSSSSVVSAAVASSSSSASMAVAAVAVSVSVAS 914  
QY 301 GACGTGACCTATGAGGCAACCTCTGTGCGGCAATGAGTGGCGGACATTACC 360  
DB 915 VVSSSSSCSSSSASVAVASVAVASVSSSSSSSVTSASVAVSVAVMSAVVSS 974  
QY 361 CCGGTGAGTCAAGAACAGGTAATCTCTGACTTTTGCACCCGCTCCGAGCGAG 420  
DB 975 ASSSSSVSVSVVAVASAAAAAASSSASAVAVASSSSSSSSSSSAS 1034  
QY 421 GAGATGACGGGAGAGTCTCGTCCGCGGAGCATCTTCGGAATCCAGA 472  
DB 1035 SVSSVSSSCSVSVSVSSSVAVASVAVASVSSVAAVAAVAAAAASA 1086

RESULT 5  
LOCUS CL945925/c 2100 bp DNA linear GSS 21-SEP-2004  
DEFINITION OIRFSB001364 Oryza sativa Expressed Library Oryza sativa (indica  
cultivar-group) genomic, genomic survey sequence.

ACCESSION CL945925.1 GI:52357934  
VERSION CL945925  
KEYWORDS GSS.  
SOURCE Oryza sativa (indica cultivar-group)  
ORGANISM Oryza sativa (indica cultivar-group)

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 2100)

REFERENCE Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,  
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,  
Wong, G. K. S., Deng, X. W. and Wang, J.  
An analysis of transcriptional regulation of the rice genome and  
its comparison to Arabidopsis  
Unpublished (2004)  
Contact: Chen Chen

COMMENT Department of Bioinformatic  
Beijing Institute of Genomics  
Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Fax: 86-10-8048676  
Email: chenchen@genomics.org.cn  
Rice genomic sequence.  
Class: exon-trapped.

## FEATURES

source

Location/Qualifiers  
1. 2100  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:39946"  
/clone\_lib="Oryza sativa Expressed Library"  
/note="Oryza sativa exon trapped genomic sequences"

## ORIGIN

Query Match 4.3%; Score 44.4; DB 9; Length 2100;  
Best Local Similarity 59.5%; Pred. No. 1.7;

Matches 75; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 639 GACAGATTCACCGTGTCTTCTGCTGCGGAGCTTGCACAAAGCTTTTGTAGCTC 698  
 174 GAGCGGAACACCGCCCGCTCGGCTCTGCGGCGGAGCCACCATGCGGTAGCGCGCC 115

DB 699 CGGACCGTGGCGCTGACAGCGGCAACCGCCCTCTGACCAATCTCCGACGA 758  
 114 CCGCGCATGCGGCTGCGGCGCGCCAGCGCATCCCGCGCGCCCGCCACACCGGCTG 55

QY 759 CGGCGG 764  
 54 CGGCGG 49

DB 54 CGGCGG 49

RESULT 6  
 BP153627/c 506 bp mRNA linear EST 30-DEC-2003  
 LOCUS BP153627 full-length enriched swine cDNA library, adult ovary Sus  
 DEFINITION BP153627  
 ACCESSION BP153627  
 VERSION BP153627  
 KEYWORDS BP153627.1 GI:40403100  
 SOURCE EST  
 ORGANISM Sus scrofa (pig)  
 Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 REFERENCES 1 (bases 1 to 506)  
 Unishii,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H.,  
 Okumura,N., Hamabima,N. and Awata,T.  
 PEDF (Pig EST Data Explorer): construction of a database for ESTs  
 derived from porcine full-length cDNA libraries  
 Nucleic Acids Res. 32 (1), D484-D488 (2004)  
 CONTACT: Hirohide Unishii  
 Address: Animal Genome Laboratory, Genome Research Department  
 National Institute of Agrobiological Sciences  
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan  
 Tel: +81-29-838-8627  
 Fax: +81-29-838-8627  
 Email: huenishi@affrc.go.jp  
 EST project with full-length enriched cDNA libraries carried out in  
 Animal Genome Research Program (Japan) by National Institute of  
 Agrobiological Sciences and STRAF-Institute  
 Single pass sequencing of clones derived from oligo-capped cDNA  
 library  
 Vector sequences were eliminated by RepeatMasker version 2002/07/13  
 and crossmatch version 0.990319  
 Low quality bases were trimmed based on the quality values.  
 Location/Qualifiers  
 1..506  
 /organism="Sus scrofa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9823"  
 /clone="OVRM10094G12"  
 /tissue\_type="ovary"  
 /dev\_stage="adult"  
 /clone\_lib="full-length enriched swine cDNA library, adult  
 ovary"

ORIGIN

Query Match 4.3%; Score 44; DB 5; Length 506;  
 Best Local Similarity 48.1%; Pred. No. 1.8;  
 Matches 125; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 539 AGGCCCCGAGTACGTCCTCCAGTGTACATGCGGCCCACTGCACCTACTAAGACCC 598  
 389 AGCGCTGGGCTGGCTCTCCAGCGCTCTTGGACCTTAATCTCATGATAAACG 330

QY 599 CTCCTATACACCTCCAGGAGAAAGAGAGATTCTTGACAGAAATCCACCGTGTCTT 658  
 329 TGTGTGTGGGCGGCAACCGGCTCCGGGAGCGAGCCGACCGCGCCCTTTCCCG 270

DB 329 TGTGTGTGGGCGGCAACCGGCTCCGGGAGCGAGCCGACCGCGCCCTTTCCCG 270

QY 659 CTGCCCTGGGAGCTTTGCCACAAAGCTTTTGTAGTCTCGGACCGTGGCGGTGACA 718

DB 269 CAGCGCAGCGCCGAGCTCCCGAGCGACGACCGACCGGCGGACCGACCGCGCCT 210  
 719 GCGGACGCGCAACCGCCCTCTGACCAATCTCGACGAGCGGCGGACAGATCTGACG 778  
 209 CGGGTCGCGACGAGGCGCTCGAGGCCCGCCACCGGAGGCGGCGGACCGGTCGCCG 150

QY 779 TTGAGTCGTATTCCTCATG 798  
 149 GGGGCGGAGGCGCACGAG 130

DB 149 GGGGCGGAGGCGCACGAG 130

RESULT 7  
 CN010969/c 652 bp mRNA linear EST 29-MAR-2004  
 LOCUS CN010969  
 DEFINITION CN010969  
 ACCESSION CN010969  
 VERSION CN010969.1 GI:45801001  
 KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Pooidae; Triticeae; Triticum.  
 REFERENCES 1 (bases 1 to 652)  
 Anderson,O.D., Chao,S., Han,P.S., Heinlen,S., Hsia,C.C., Kang,Y.,  
 Krueger,W.M., Izzo,G.R., Miller,S., Muehlbauer,G.J., Miller,R.,  
 Pritsch,C., Rauech,C.U., Seaton,C.L., Tong,J.C., Vance,C. and  
 Wilson,C.F.  
 The structure and function of the expressed portion of the wheat  
 genomes - Fusarium graminearum infected spike cDNA library  
 Unpublished (2001)  
 CONTACT: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818  
 Email: oanderson@nwg.ars.gov  
 Sequences have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20. No effort was taken  
 to identify ESTs of fungal origin from this library, thus this EST  
 could be of wheat or fungal origin.  
 Seq primer: SK primer.  
 Location/Qualifiers  
 1..652  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Suma13"  
 /db\_xref="taxon:4565"  
 /clone="WHE3878\_D02\_H04"  
 /tissue\_type="Spike"  
 /dev\_stage="Adult plant"  
 /lab\_host="E. coli SOLR"  
 /clone\_lib="wheat Fusarium graminearum infected spike cDNA  
 library"

ORIGIN

/note="Vector: Lambda Uni-ZAP XR, excised phagemid  
 pluescript SK; Site\_1: EcoRI; Site\_2: XhoI; Plants were  
 grown in the greenhouse. Spikes were sprayed at anthesis  
 with Fusarium graminearum. Total RNA, and poly(A) RNA were  
 prepared and pooled from infected spike at 0, 6, 12, 24,  
 36 and 48 hours after inoculation. a cDNA library was  
 made, and the cDNA clones were in vivo excised to give  
 pluescript phagemids in G. Muehlbauer lab at the  
 University of Minnesota (Krueger, W.M., Muehlbauer, G.J.,  
 Pritsch, C., Vance, C.). The cDNA library should contain  
 genes of both wheat and fungal pathogen origin. Plasmid  
 DNA preparations and DNA sequencing were performed in the  
 OD Anderson lab (all other authors)."

Query Match 4.3%; Score 44; DB 7; Length 652;  
 Best Local Similarity 48.8%; Pred. No. 1.8; Indels 125; Gaps 0;  
 Matches 119; Conservative 0; Mismatches 0; Gaps 0;

QY 662 CCCTGGCGAGCTTGGCCACAAAGGCTTTTGTAGCTCCGACCGTGGCCGTGAGACG 721  
 DB 270 CCGTCCCGGAGAGAGCCGAGTACATCCGCGCCGCGCTGAGCTCCGCGCTTTTCCG 211

QY 722 GCACGGCAACCGCCCTCTCTGACCAATCTCTCCAGCAAGCGGAGACAGATCTGAGCTTG 781  
 DB 210 CCGTGGCCACCGCCACCGCCACCGTACCTCTCCGCGCGCTGAGAGCCGCGCGCCG 151

QY 782 AGTCGATTTCTTCATGCCCCCTTGAAGGAGAGCGCGGAGACCCCGATCTGAGGAG 841  
 DB 150 TAACTTCACCGCTTCCGACCGTGGCCAGGAGTACCGCCGCGCGCTGACCTTCCA 91

QY 842 GGTCTTGCTTACCTGAGTGAAGAGCGCGTGAAGACGTCTGCTGCTGATGCTCT 901  
 DB 90 CCGCTTCCCGACCGTGGCGAGGTAAGCGCCGACCGCGCATAGCTCTCCGCGCTTCC 31

QY 902 ACAC 905  
 DB 30 CCAC 27

RESULT 8  
 CA679773 587 bp mRNA linear EST 24-NOV-2002  
 LOCUS wlm4.pk0024.c7 wlm4 Triticum aestivum cDNA clone wlm4.pk0024.c7 5'  
 DEFINITION end, mRNA sequence.  
 ACCESSION CA679773  
 VERSION CA679773.1 GI:25264502  
 KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.  
 REFERENCE 1 (bases 1 to 587)  
 Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,  
 Miao, G., Caraher, N. and Hanafey, M.K.  
 Dupont Wheat cDNA Sequence  
 Unpublished (2002)  
 CONTACT: Scott V. Tingey  
 CROP GENETICS  
 E. I. Dupont de Nemours and Company  
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA  
 Tel: 302-631-2602  
 Fax: 302-631-2607  
 Email: Scott.V.Tingey@USA.dupont.com  
 Seq primer: M13.  
 Location/Qualifiers  
 1..587  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Stephens"  
 /db\_xref="taxon:4565"  
 /clone="wlm4.pk0024.c7"  
 /cissue\_type="leaf"  
 /clone\_lib="wlm4"  
 /note="Vector: pBluescript SK+; Site\_1: EcoRI; Site\_2:  
 XhoI; Wheat (Triticum aestivum L.) seedlings 4 hr after  
 inoculation w/E. graminis"

ORIGIN  
 Query Match 4.2%; Score 43.4; DB 6; Length 587;  
 Best Local Similarity 53.3%; Pred. No. 2.6; Indels 78; Gaps 0;  
 Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 445 CCGGCGAAGTCTTGGGAAATCAAGAAATTTCCACGAGATGCCCGATGAGGAGCG 504  
 DB 217 CCGCGCGCGCTGCTCCGAAACCTTAATCTCGCTCCGCGCGCGATCTCTCCGCG 276

QY 505 CCGATTACCAACCCCTCGCTGAGTCTGTGAAGCCCGGACTACGCTCCGAGTG 564  
 DB 277 CCGGCGCTCCGTATGCTCCGCTCTTGTATTTCCGGAAGCCGCTTCTTCCGCGCAAC 336

QY 565 GTACATGAGGTGCCACTGCGACCTACTTAAGACCCCTCTAATACACC 611  
 DB 337 TGGCAAGACCAACGCTCTTAATTATAGCCGCAATACATCAAC 383

RESULT 9  
 CV039737 487 bp mRNA linear EST 01-OCT-2004  
 LOCUS 3AL.25G10.5', mRNA sequence.  
 DEFINITION 3AL.25G10.5', mRNA sequence.  
 ACCESSION CV039737  
 VERSION CV039737.1 GI:53559176  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 487)  
 Evock-Clover, C.M., Ashwell, C.M., McMurtry, J.P., Lillehoj, H.S.,  
 Matukumalli, L.K. and Van Tassell, C.P.  
 Characterization of expressed sequence tags generated from multiple  
 chicken tissues  
 Unpublished (2004)  
 CONTACT: Christina M. Clover  
 Growth Biology Laboratory  
 Animal and Natural Resources Institute  
 Bldg. 200 Rm215 BARC-East, Beltsville, MD 20705, USA  
 Tel: 3015048224  
 Fax: 3015048623  
 Email: chrisce@mri.barc.usda.gov  
 Single pass sequencing. Bases called and trimmed with phred  
 0.000925 using options -trim alt '- -trim fastavector identified by  
 cross\_match using options -mismatch 12 -minscore 12  
 Place: 25 row: G column: 10  
 Seq primer: CCCAGTCACGACGTGTAAACG  
 High quality sequence stop: 487.  
 Location/Qualifiers  
 1..487  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Leghorn and broiler"  
 /db\_xref="taxon:9031"  
 /clone="3AL.25G10"  
 /lab\_host="DH5alpha"  
 /clone\_lib="BARC\_3GAL chicken mixed tissue"  
 /note="Vector: pBluescript, SK+, Stratagene, Site\_1: NotI;  
 Site\_2: EcoRI; Normalized library of pooled RNA isolated  
 from whole brain, ultimobranchial gland, parathyroid  
 gland, cecal tonsil and primordial germ cells Multiple"

ORIGIN  
 Query Match 4.2%; Score 43; DB 7; Length 487;  
 Best Local Similarity 55.8%; Pred. No. 3.3; Indels 65; Gaps 0;  
 Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 318 GGGCAACCTCTGTGGCGGATGAGATGGCGGAGCATTTACCCGCTGAGACAGAA 377  
 DB 82 GGGGACCGCTGACCGGAGGTGAGGCGCGCGCTGACCGGAGAGGCTTACG 141

QY 378 CAAGTATATCTTGACTCTTTCGACCCGCTCCAGCGGAGAGATGAGCGGAGAT 437  
 DB 142 AAGAGTGGATCCTTAACCGCGCATTCAGCTGTGTAAGAGAGAGCTGACCGGCGCA 201

QY 438 GTCCGTCGCGGAGAGATCTTGGGAA 464  
 DB 202 GGAGCGCTTCGCGACCGCTTTCAGAA 228

RESULT 10  
CD16293  
LOCUS  
DEFINITION  
4031393 16L - Chicken Intestinal Lymphocyte Gallus gallus cDNA  
ACCESSION  
CD16293  
VERSION  
CD16293.1 GI:32277947  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Gallus gallus (chicken)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE  
1 (bases 1 to 540)  
Min, W., Lillehoj, H.S., Ashwell, C.M., Matukumalli, L.K., van  
Tassel, C. and Han, J.Y.  
Chicken Intestinal Lymphocyte EST database as a resource for the  
analysis of mucosal immune function  
Unpublished (2003)  
CONTACT: Hyun S. Lillehoj  
Animal Parasite Diseases Laboratory  
Animal and Natural Resources Institute, USDA  
Bldg. 1043, BARC-East, Beltsville, MD 20705, USA  
Tel: 3015048771  
Fax: 3015045103  
Email: lilleho@anri.barc.usda.gov  
Single pass sequencing. Bases called and trimmed with phred  
0.000925 using options -trim\_alt '' -trim\_fasta. Vector identified  
by cross\_match using options -mismatch 12 -mismatch 12  
Plate: 60 row: N column: 04  
Seq primer: ATTTAGTGACACTATAG  
High quality sequence stop: 540.  
Location/Qualifiers  
1..540  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="white leghorn SC"  
/db\_xref="taxon:9031"  
/clone="16L 60N04"  
/sex="mixed"  
/tissue\_type="Gut"  
/cell\_type="Lymphocyte"  
/dev\_stage="Adult"  
/lab\_host="EMD10B"  
/clone\_lib="16L - Chicken Intestinal Lymphocyte"  
/note="Organ: Intestine; Vector: pCMV-SPORT6; Site: 1;  
Salt: Site 2: Not; Normalized library from chicken gut  
infected with coccidia duodenum and middle gut."

FEATURES  
source  
1..540  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="white leghorn SC"  
/db\_xref="taxon:9031"  
/clone="16L 60N04"  
/sex="mixed"  
/tissue\_type="Gut"  
/cell\_type="Lymphocyte"  
/dev\_stage="Adult"  
/lab\_host="EMD10B"  
/clone\_lib="16L - Chicken Intestinal Lymphocyte"  
/note="Organ: Intestine; Vector: pCMV-SPORT6; Site: 1;  
Salt: Site 2: Not; Normalized library from chicken gut  
infected with coccidia duodenum and middle gut."

Query Match  
Best Local Similarity 55.8%; Pred. No. 3.3;  
Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 318 GGCACCTCTCTGTGGCGCATGATGGCGGGGACATTACCCGCTGAGTCAGAGAA 377  
DB 99 GCGGACCGCCTGACGCGGAGGTGAGGCGGCCCTTAACCGGACACAGCTGAGGC 158  
QY 378 CAAGTAGTAATCTTGACTCTTTTGACCCGCTCCGAGCGAGAGAGATGACGGGAAGT 437  
DB 159 AGAGGGGATCCCTGTAACCGCGCATCTCAAGCTGTGAGAGAGAGCTGACCGGCGCCA 218  
QY 438 GTCCGTCGCGGCGAGATCCTGCGGAA 464  
DB 219 GAGGCGCTCGCACCGCTCTGAGAA 245

RESULT 11  
CD16293  
LOCUS  
DEFINITION  
CD16293 573 bp mRNA linear EST 20-MAY-2003  
ppgzn.pk010.j13 Normalized chicken pituitary/hypothalamus/pineal  
cDNA library (ppgzn) Gallus gallus cDNA clone ppgzn.pk010.j13 5'  
similar to pitr167850|167850 tropomyosin - rat >gb|AAA42264.1|

ACCESSION  
CD16293  
VERSION  
CD16293.1 GI:30956292  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Gallus gallus (chicken)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE  
1 (bases 1 to 573)  
Porter, T.E. and Cogburn, L.A.  
Chicken ESTs from pituitary/hypothalamus/pineal gland  
Unpublished (2001)  
CONTACT: Larry A. Cogburn  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1335  
Fax: 302-831-2822  
Email: cogburn@udel.edu, www.chickest.udel.edu.  
Location/Qualifiers  
1..573  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="Commercial broiler chickens"  
/db\_xref="taxon:9031"  
/clone="ppgzn.pk010.j13"  
/sex="Male and Female"  
/tissue\_type="Pituitary Gland/Hypothalamus/Pineal Gland"  
/dev\_stage="Embryonic (d12,d14,d19); post-hatch (1,3,5,7,9  
weeks)"  
/lab\_host="E. Coli EMD10B"  
/clone\_lib="Normalized chicken  
pituitary/hypothalamus/pineal cDNA library (ppgzn)"  
/note="Vector: pCMVSPORT6; Library made from equivalent  
pools of total RNA isolated from each tissue at different  
ages. Single pass sequencing from 5'-end"

FEATURES  
source  
1..573  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="Commercial broiler chickens"  
/db\_xref="taxon:9031"  
/clone="ppgzn.pk010.j13"  
/sex="Male and Female"  
/tissue\_type="Pituitary Gland/Hypothalamus/Pineal Gland"  
/dev\_stage="Embryonic (d12,d14,d19); post-hatch (1,3,5,7,9  
weeks)"  
/lab\_host="E. Coli EMD10B"  
/clone\_lib="Normalized chicken  
pituitary/hypothalamus/pineal cDNA library (ppgzn)"  
/note="Vector: pCMVSPORT6; Library made from equivalent  
pools of total RNA isolated from each tissue at different  
ages. Single pass sequencing from 5'-end"

Query Match  
Best Local Similarity 55.8%; Pred. No. 3.3;  
Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 318 GGCACCTCTCTGTGGCGCATGATGGCGGGGACATTACCCGCTGAGTCAGAGAA 377  
DB 103 GCGGACCGCCTGACGCGGAGGTGAGGCGGCCCTTAACCGGACACAGCTGAGGC 162  
QY 378 CAAGTAGTAATCTTGACTCTTTTGACCCGCTCCGAGCGAGAGAGATGACGGGAAGT 437  
DB 163 AGAGGTGATCCCTGTAACCGCGCATCTCAAGCTGTGAGAGAGAGCTGACCGGCGCCA 222  
QY 438 GTCCGTCGCGGCGAGATCCTGCGGAA 464  
DB 223 GAGGCGCTCGCACCGCTCTGAGAA 249

RESULT 12  
B1391789  
LOCUS  
DEFINITION  
B1391789 591 bp mRNA linear EST 06-AUG-2001  
ppgzn.pk005.i7 Normalized Chicken Pituitary/Hypothalamus/Pineal  
Library Gallus gallus cDNA clone ppgzn.pk005.i7 5' similar to  
pitr153784|153784 tropomyosin - rat >gb|AAA21721.1| (U24775)  
tropomyosin [Rattus norvegicus], mRNA sequence.

ACCESSION  
B1391789  
VERSION  
B1391789.1 GI:15085071  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Gallus gallus (chicken)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE  
1 (bases 1 to 591)  
Porter, T.E. and Cogburn, L.A.  
ESTs from Normalized Chicken Pituitary/Hypothalamus/Pineal cDNA

JOURNAL  
COMMENT

Library USDA/IFAFS Animal Genome Project  
Unpublished (2001)  
Contact: Larry A. Cogburn  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1335  
Fax: 302-831-2822  
Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES  
source

Location/Qualifiers

1. .591  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="Commercial broiler chicken"  
/db\_xref="taxon:9031"  
/clone="pgfp1.pk001.f7"  
/sex="Male and Female"  
/tissue\_type="Pituitary Gland/Hypothalamus/Pineal Gland"  
/dev\_stage="Embryonic (dl2,dl4,dl9); post-hatch  
(w1,w3,w5,w7,w9)"  
/lab\_host="E. coli EMDH10B"  
/clone\_lib="Normalized Chicken  
Pituitary/Hypothalamus/Pineal Library"  
/note="Vector: PCWVSPORT6; Library made from equivalent  
pools of total RNA isolated from each tissue at different  
ages. Single pass sequencing from 5'-end"

## ORIGIN

## Query Match

Best Local Similarity 55.8%; Pred. No. 3.3;  
Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 318 GGCACACCTCTCTGCGGCGATGAGTGGCGGAGCATTAACCGCGTGAAGTCAGAGAA 377  
DB 146 GGGGACCCCGCTGCACCGGAGGTGAGCGCCGCGTAAACCGAGCAGGCTGAGGC 205  
QY 378 CAAGTAGTATCTCTGACTCTTTGACCGCGCTCCGAGCGAGAGATGAGCGGAGAT 437  
DB 206 AGAGGTGGATCTCTTAACCGCGCATTCAGCTGTGTAAGAGAGAGAGCTGAGCGGAGCCCA 265  
QY 438 GTCCGTCCCGCGGAGATCTCTGCGAA 464  
DB 266 GGAGCGCTCTGCACCGCTCTGAGAA 292

RESULT 13  
LOCUS

BM425602 601 bp mRNA linear EST 30-JAN-2002  
pgf2c.pk001.e8 Primary Chicken Abdominal Fat Library (pgf2c) Gallus  
gallus CDNA clone pgf2c.pk001.e8 5' similar to p1r1167849|167849

## DEFINITION

tropomyosin - rat gbl|AA42263.1| (U24776) tropomyosin [Rattus  
norvegicus], mRNA sequence.

## ACCESSION

BM425602  
EST.

## VERSION

BM425602.1 GI:18429216

## KEYWORDS

## SOURCE

## ORGANISM

Gallus gallus (chicken)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 601)  
ESTs from Primary Chicken Fat cDNA library-USDA/IFAFS Animal Genome  
Project

JOURNAL  
COMMENT

Unpublished (2002)  
Contact: Larry A. Cogburn  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1335  
Fax: 302-831-2822  
Email: cogburn@udel.edu, www.chickest.udel.edu.

## FEATURES

## source

Location/Qualifiers  
1. .601  
/organism="Gallus gallus"

/mol\_type="mRNA"  
/strain="Commercial broiler, Ottawa Research Centre,  
leghorn"

/db\_xref="taxon:9031"  
/clone="pgf2c.pk001.e8"  
/sex="Male and Female"  
/tissue\_type="Abdominal Fat"  
/dev\_stage="Embryonic (dl8,dl9); post-hatch  
(dl,w1,w3,w5,w7,w9,w16,1yr)"  
/lab\_host="E. coli EMDH10B"  
/clone\_lib="Primary Chicken Abdominal Fat Library (pgf2c)"  
/note="Vector: PCWVSPORT6; Library made from equivalent  
pools of total RNA isolated from each developmental age  
(across strains); Single pass sequencing from 5'-end"

## ORIGIN

## Query Match

Best Local Similarity 4.2%; Score 43; DB 4; Length 601;  
Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 318 GGCACACCTCTCTGCGGCGATGAGTGGCGGAGCATTAACCGCGTGAAGTCAGAGAA 377  
DB 50 GGGGACCCCGCTGCACCGGAGGTGAGCGCCGCGTAAACCGAGCAGGCTGAGGC 109  
QY 378 CAAGTAGTATCTCTGACTCTTTGACCGCGCTCCGAGCGAGAGATGAGCGGAGAT 437  
DB 110 AGAGGTGGATCTCTTAACCGCGCATTCAGCTGTGTAAGAGAGAGAGCTGAGCGGAGCCCA 169  
QY 438 GTCCGTCCCGCGGAGATCTCTGCGAA 464  
DB 170 GGAGCGCTCTGCACCGCTCTGAGAA 196

RESULT 14  
LOCUS

BM425623 615 bp mRNA linear EST 30-JAN-2002  
pgf2c.pk001.f7 Primary Chicken Abdominal Fat Library (pgf2c) Gallus  
gallus CDNA clone pgf2c.pk001.f7 5' similar to p1r1167849|167849

## DEFINITION

tropomyosin - rat gbl|AA42263.1| (U24776) tropomyosin [Rattus  
norvegicus], mRNA sequence.

## ACCESSION

BM425623  
EST.

## VERSION

BM425623.1 GI:18429255

## KEYWORDS

## SOURCE

## ORGANISM

Gallus gallus (chicken)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 615)  
ESTs from Primary Chicken Fat cDNA library-USDA/IFAFS Animal Genome  
Project

JOURNAL  
COMMENT

Unpublished (2002)  
Contact: Larry A. Cogburn  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1335  
Fax: 302-831-2822  
Email: cogburn@udel.edu, www.chickest.udel.edu.

## FEATURES

## source

Location/Qualifiers  
1. .615  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="Commercial broiler, Ottawa Research Centre,  
leghorn"  
/db\_xref="taxon:9031"  
/clone="pgf2c.pk001.f7"  
/sex="Male and Female"  
/tissue\_type="Abdominal Fat"  
/dev\_stage="Embryonic (dl8,dl9); post-hatch  
(dl,w1,w3,w5,w7,w9,w16,1yr)"  
/lab\_host="E. coli EMDH10B"  
/clone\_lib="Primary Chicken Abdominal Fat Library (pgf2c)"

/note="Vector: PCMVSPORT6; Library made from equivalent pools of total RNA isolated from each developmental age (across strains); Single pass sequencing from 5'-end"

Job time : 3613.18 secs

## ORIGIN

Query Match 4.2%; Score 43; DB 4; Length 615;  
Best Local Similarity 55.8%; Pred. No. 3.3;  
Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 318 GGGCAACTCTGTCGGCCGATGAGATGGCGGACATTACCGCGTGAATCAGAGAA 377  
DB 104 GCGGACCGCGCTGACGAGGTGAGCGCGCGCTAACCGCGACAGCTGAGGC 163  
QY 378 CAAGTAGTAATCTGACTCTTTGACCGCGCTCCGAGGAGAGATGACGGGAAGT 437  
DB 164 AGAGGTGACATCCCTGAAACCGCGCATCAGCTGTAAGAGAGAGCTGACCGGGCCA 223  
QY 438 GTCCGTCCCGCGGAGATCTCGGAA 464  
DB 224 GGAGCGCTCGCCACCGCTGCAAGAA 250

## RESULT 15

LOCUS AJ450428 664 bp mRNA linear EST 19-APR-2002  
DEFINITION AJ450428 riken1 Gallus gallus CDNA clone 25d13r1, mRNA sequence.  
ACCESSION AJ450428  
VERSION AJ450428.1 GI:20217649  
KEYWORDS EST  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 664)  
Buerstedde,J.M.  
Gallus gallus bursal lymphocyte EST  
Unpublished (2002)  
Contact: Buerstedde JM  
Cellular Immunology  
Heinrich-Pette-Institute  
Martinistr. 52, 20251 Hamburg, Germany  
Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.  
Location/Qualifiers

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

## FEATURES

source 1..664  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone="25d13r1"  
/cell\_type="bursal lymphocyte"  
/dev\_stage="2-3 weeks old"  
/clone\_id="riken1"  
/note="CB indred strain"

## ORIGIN

Query Match 4.2%; Score 43; DB 1; Length 664;  
Best Local Similarity 55.8%; Pred. No. 3.4;  
Matches 82; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 318 GGGCAACTCTGTCGGCCGATGAGATGGCGGACATTACCGCGTGAATCAGAGAA 377  
DB 96 GCGGACCGCGCTGACGAGGTGAGCGCGCGCTAACCGCGACAGCTGAGGC 155  
QY 378 CAAGTAGTAATCTGACTCTTTGACCGCGCTCCGAGGAGAGATGACGGGAAGT 437  
DB 156 AGAGGTGACATCCCTGAAACCGCGCATCAGCTGTAAGAGAGAGCTGACCGGGCCA 215  
QY 438 GTCCGTCCCGCGGAGATCTCGGAA 464  
DB 216 GGAGCGCTCGCCACCGCTGCAAGAA 242

Search completed: February 19, 2005, 11:38:15





# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 144962**

**TO: Bao-Qun Li**  
**Location: rem/3d24/3c18**  
**Art Unit: 1648**  
**Tuesday, February 22, 2005**

**Case Serial Number: 09/664363**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**REM-1A55**  
**Phone: 571-272-2512**

**edward.hart@uspto.gov**

### **Search Notes**

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**This Page Blank (uspto)**

STIC-Biotech/ChemLib

144962

mg

From: Li, Bao-Qun  
Sent: Friday, February 11, 2005 8:08 AM  
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF SEQ ID NO; 3 IN APPLICATION SN.  
09,664,363. THANKS.

Bao Qun Li M.D  
TC 1600  
Art Unit 1648  
Tel. 517-272-0904  
REM, 3C18  
Rm. 3D24

RECEIVED  
FEB 11 2005  
STIC

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: 2/16/05  
Date Completed: 2/22/05  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA Sequence: # 1  
AA Sequence: # \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: PSA  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**



# STIC SEARCH RESULTS

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 21:00:40 ; Search time 7668.71 Seconds  
(without alignments)  
10899.589 Million cell updates/sec

Title: US-09-664-363-3

Perfect score: 1770  
Sequence: 1 CAAATGACTTCCAGACCC.....ACTACTCACCCTGACCG 1770

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ha:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sbs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1770	100.0	1770	6 A28131	A28131 PT-NANBH mr
2	1770	100.0	1770	6 A32184	A32184 NANBH PT p
3	1770	100.0	1770	6 ARI44032	ARI44032 Sequence
4	1763.6	99.6	3750	6 A28157	A28157 PT-NANBH mr
5	1763.6	99.6	3750	6 A32203	A32203 NANBH PT p
6	1763.6	99.6	3750	6 ARI44051	ARI44051 Sequence
7	1580.6	89.3	9373	14 ARI13916	ARI13916 Hepatitis
8	1570.6	88.7	9373	14 ARI208024	ARI208024 Hepatitis
9	1562.6	88.3	9373	14 ARI587845	ARI587845 Hepatitis
10	1562.6	88.3	9410	14 ARI587845	ARI587845 Hepatitis
11	1562.6	88.3	9547	14 ARI587845	ARI587845 Hepatitis
12	1557.8	88.0	9374	14 ARI587845	ARI587845 Hepatitis
13	1557	88.0	9425	14 ARI587845	ARI587845 Hepatitis
14	1557	88.0	9573	14 ARI587845	ARI587845 Hepatitis
15	1556.2	87.9	9410	14 ARI587845	ARI587845 Hepatitis
16	1554.6	87.8	9379	14 ARI587845	ARI587845 Hepatitis
17	1553.8	87.8	8638	6 ARI587845	ARI587845 Hepatitis
18	1553.8	87.8	8638	6 ARI587845	ARI587845 Hepatitis
19	1553.8	87.8	8638	6 ARI587845	ARI587845 Hepatitis

20	1553.8	87.8	8638	6 AX472315	AX472315 Sequence
21	1553.4	87.8	8642	6 ARI587845	ARI587845 Hepatitis
22	1553.4	87.8	8642	6 ARI587845	ARI587845 Hepatitis
23	1552.2	87.7	7989	6 ARI587845	ARI587845 Hepatitis
24	1552.2	87.7	7989	6 ARI587845	ARI587845 Hepatitis
25	1552.2	87.7	7989	6 ARI587845	ARI587845 Hepatitis
26	1552.2	87.7	8001	6 ARI587845	ARI587845 Hepatitis
27	1552.2	87.7	8001	6 ARI587845	ARI587845 Hepatitis
28	1552.2	87.7	8001	6 ARI587845	ARI587845 Hepatitis
29	1552.2	87.7	8637	6 ARI587845	ARI587845 Hepatitis
30	1552.2	87.7	8637	6 ARI587845	ARI587845 Hepatitis
31	1552.2	87.7	8637	6 ARI587845	ARI587845 Hepatitis
32	1552.2	87.7	8637	6 ARI587845	ARI587845 Hepatitis
33	1552.2	87.7	8638	6 ARI587845	ARI587845 Hepatitis
34	1552.2	87.7	8638	6 ARI587845	ARI587845 Hepatitis
35	1552.2	87.7	8638	6 ARI587845	ARI587845 Hepatitis
36	1552.2	87.7	8639	6 ARI587845	ARI587845 Hepatitis
37	1552.2	87.7	8639	6 ARI587845	ARI587845 Hepatitis
38	1552.2	87.7	8643	6 ARI587845	ARI587845 Hepatitis
39	1552.2	87.7	8643	6 ARI587845	ARI587845 Hepatitis
40	1552.2	87.7	8649	6 ARI587845	ARI587845 Hepatitis
41	1552.2	87.7	8649	6 ARI587845	ARI587845 Hepatitis
42	1552.2	87.7	8649	6 ARI587845	ARI587845 Hepatitis
43	1552.2	87.7	9033	14 ARI587845	ARI587845 Hepatitis
44	1552.2	87.7	9410	14 ARI587845	ARI587845 Hepatitis
45	1552.2	87.7	9605	6 ARI587845	ARI587845 Hepatitis

# ALIGNMENTS

## RESULT 1

LOCUS A28131 1770 bp DNA linear PAT 07-JUN-1995  
DEFINITION PT-NANBH mRNA fragment from patent GB2239245.  
ACCESSION A28131  
VERSION A28131.1 GI:1248616  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 1770)  
AUTHORS  
TITLES Post-transfomational non-A non-B hepatitis viral polypeptides  
JOURNAL Patent: GB 2239245-A 3 26-JUN-1991;  
FEATURES  
source Location/Qualifiers  
1..1770  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/clone="UG2"  
<1..>1770  
/cramel\_start=1  
/cramel\_table=1  
/product="PT-NANBH polypeptide"  
/protein\_id="CAA01925.1"  
/db\_xref="GI:1248617"  
/translation="QNDPFDADLIEANILMRHEMGDITRVESENKVTILDSFDPIRA  
EEDEREVSVPAILRKSKFPAMPAMPADYNPPLBEMKAPDYVPVHGFPLPT  
KTPPIPPRRKRTVITESTVSALEATKFGSGFVAVSGMTAGALITPCAA  
ACSDVESVSMPLRGEPPDPLSGSMSTVEEEDVDCSGMTAGALITPCAA  
ESKUPINLNSILRHHNMVYATTSRSASQOKVTPTPLDITDHYDVKEMAK  
ASTVAKLISVEBACKLTTPHSKSKFGGADVRNLKSKALNHRISVEDLLEDET  
PIDITIMANVEFCVQPERGKRPALIVFPLGVAVCKMALYDVSTLPOAVGSS  
YGFQVSGQVREFFLVAMSKTTPMGFAVDFPCFSTVTENIRVEESLYCCDLAP  
ARQALISLTERLYIGPLTNSKONGVRCRAGVLTSCNTLTCYKAAACRAA  
KIDQCTMIVCGGLVITCSAGTQEDASLIRVFTAMTRYSPGPPQPEVDLELIT  
SCSSNVSVADHDSKRVYTLTRDP"

## CDS

ORIGIN  
Query Match 100.0%; Score 1770; DB 6; Length 1770;  
Best local similarity 100.0%; Pred. No. 0;  
Matches 1770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGGCGGATGAGATG 60  
Db 1 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGTGGCGGATGAGATG 60  
QY 61 GGGGGGGAATTACCCGCGTGAAGTACAGAACAGGTAGTAACTCTGACCTTTTCGAC 120  
Db 61 GGGGGGGAATTACCCGCGTGAAGTACAGAACAGGTAGTAACTCTGACCTTTTCGAC 120  
QY 121 CCGCTCCAGCGGAGGAGTGAAGGTGCTCCGCGGAGATCCGCGGAAA 180  
Db 121 CCGCTCCAGCGGAGGAGTGAAGGTGCTCCGCGGAGATCCGCGGAAA 180  
QY 181 TCCAAAGAAATTCACACAGCGATGCCCCGATGGGCAAGCCGATTAACCTCCGCTG 240  
Db 181 TCCAAAGAAATTCACACAGCGATGCCCCGATGGGCAAGCCGATTAACCTCCGCTG 240  
QY 241 CTGAGACTCTGGAAGGCCCCGGACTAGCTCCTCCAGTGTATCATGGTGGCCACTGCA 300  
Db 241 CTGAGACTCTGGAAGGCCCCGGACTAGCTCCTCCAGTGTATCATGGTGGCCACTGCA 300  
QY 301 CCTACTAAGACCCCTCTATACACCTCCAGGAGAAAGAGACAGTTGTCTGACAGAA 360  
Db 301 CCTACTAAGACCCCTCTATACACCTCCAGGAGAAAGAGACAGTTGTCTGACAGAA 360  
QY 361 TCCACCGTGTCTTCTGCGCGAGCTTGGCCAAAGGCTTTTGTGATCCGACCG 420  
Db 361 TCCACCGTGTCTTCTGCGCGAGCTTGGCCAAAGGCTTTTGTGATCCGACCG 420  
QY 421 TGGCGCGTGAAGCGGCAAGCGACCCGCTCTGACCAATCTCTGCAAGCGCGGA 480  
Db 421 TGGCGCGTGAAGCGGCAAGCGACCCGCTCTGACCAATCTCTGCAAGCGCGGA 480  
QY 481 GCAGAGCTGACGTTGAGTGTATTCCTCATGCCCCCTTGAGGGGAGCGGGGAGAC 540  
Db 481 GCAGAGCTGACGTTGAGTGTATTCCTCATGCCCCCTTGAGGGGAGCGGGGAGAC 540  
QY 541 CCGGATCTCAGCGACGGGTCTTGCTTACCGTGAAGTGAAGAGCGCGTGAAGACGTGTC 600  
Db 541 CCGGATCTCAGCGACGGGTCTTGCTTACCGTGAAGTGAAGAGCGCGTGAAGACGTGTC 600  
QY 601 TGTGCTGATGTCTTCAACATGAGAGGCGCTGTATATACCGCATGCGCTGAGAGAA 660  
Db 601 TGTGCTGATGTCTTCAACATGAGAGGCGCTGTATATACCGCATGCGCTGAGAGAA 660  
QY 661 AGCAAGCTGCCATCAACCGGTTGAGCAACTCTTGCTGCTGATCCACAACATGGTCTAC 720  
Db 661 AGCAAGCTGCCATCAACCGGTTGAGCAACTCTTGCTGCTGATCCACAACATGGTCTAC 720  
QY 721 GCTACCAATCCCGACGCGCAAGCCAGCGGAGAAAGAGTCACTTTGACAGACTGCA 780  
Db 721 GCTACCAATCCCGACGCGCAAGCCAGCGGAGAAAGAGTCACTTTGACAGACTGCA 780  
QY 781 ATCTTGAGATGATCACTACAGACGCTGCTCAAGAGATGAAGCGGAGCGCTGCACATT 840  
Db 781 ATCTTGAGATGATCACTACAGACGCTGCTCAAGAGATGAAGCGGAGCGCTGCACATT 840  
QY 841 AAGGCTAGCTTCTATCAGTAGAGAAAGCTGCAAGCTGACGCCCCCACTTCGCGCAA 900  
Db 841 AAGGCTAGCTTCTATCAGTAGAGAAAGCTGCAAGCTGACGCCCCCACTTCGCGCAA 900  
QY 901 TCTAAATTTGGCTATGAGGCAAGAGCTCCGGAACCTATCCAGAGGCGCATTTAACAC 960  
Db 901 TCTAAATTTGGCTATGAGGCAAGAGCTCCGGAACCTATCCAGAGGCGCATTTAACAC 960  
QY 961 ATCCGCTCGTGTGGAGAGACTTGTGGAAGACCTGAAACCAATTGACACCACTATC 1020  
Db 961 ATCCGCTCGTGTGGAGAGACTTGTGGAAGACCTGAAACCAATTGACACCACTATC 1020  
QY 1021 ATGGCAAAATGAGGTTTCTGCGTCAACCAAGAGAGAGGCGCGCAAGCGAGCTGCG 1080  
Db 1021 ATGGCAAAATGAGGTTTCTGCGTCAACCAAGAGAGAGGCGCGCAAGCGAGCTGCG 1080

QY 1081 CTTATGATGTTCCAGACTTGGGGGTCCGTGTGTGTCGAGAAAATGCGCCCTATGACGTG 1140  
Db 1081 CTTATGATGTTCCAGACTTGGGGGTCCGTGTGTGTCGAGAAAATGCGCCCTATGACGTG 1140  
QY 1141 GTCTCACCTTCCTCAAGCTGTGATGGGCTTCCTGTACGGAATTCAGTATTTCTCTGGA 1200  
Db 1141 GTCTCACCTTCCTCAAGCTGTGATGGGCTTCCTGTACGGAATTCAGTATTTCTCTGGA 1200  
QY 1201 CAGCGGATCGAGTTCTGCTGTGAACGCTGGAATAATCAAGAAAGACCCCTATGGGCTTTGCA 1260  
Db 1201 CAGCGGATCGAGTTCTGCTGTGAACGCTGGAATAATCAAGAAAGACCCCTATGGGCTTTGCA 1260  
QY 1261 TATGACACCGCTGTTTGACTCAACAGTCACTGAAATGACATCCGTGAGAGATCA 1320  
Db 1261 TATGACACCGCTGTTTGACTCAACAGTCACTGAAATGACATCCGTGAGAGATCA 1320  
QY 1321 ATTTATCAATGTTGACTTGGCCCCGGAAGCCAGACAGGCGCATAGGTGCTCAGAG 1380  
Db 1321 ATTTATCAATGTTGACTTGGCCCCGGAAGCCAGACAGGCGCATAGGTGCTCAGAG 1380  
QY 1381 CCGCTTATCGGGGCTCCCTGACTAATTCAAAAGGGGAGAACTGGGCTATGCGCG 1440  
Db 1381 CCGCTTATCGGGGCTCCCTGACTAATTCAAAAGGGGAGAACTGGGCTATGCGCG 1440  
QY 1441 TCCCGCGGAGCGGCTGTGACGACTAGCTGCGGTATACCTTCACTGTTACTTGAAG 1500  
Db 1441 TCCCGCGGAGCGGCTGTGACGACTAGCTGCGGTATACCTTCACTGTTACTTGAAG 1500  
QY 1501 GCTTGTGAGCTGTGAGCTGCAAAAGCTCCAGACCTGCAAGATGCTGTGTGCGGAGAC 1560  
Db 1501 GCTTGTGAGCTGTGAGCTGCAAAAGCTCCAGACCTGCAAGATGCTGTGTGCGGAGAC 1560  
QY 1561 GGCCTTGTGTTATCTGAGAGCGCGGGAACCCAGAGAGACGCGGCGAGCTACGAGTC 1620  
Db 1561 GGCCTTGTGTTATCTGAGAGCGCGGGAACCCAGAGAGACGCGGCGAGCTACGAGTC 1620  
QY 1621 TTCAAGAGGCTATGACTAGGTACTTGCCCCCCCCGGGAGACCCGCCCAACAGAAATAC 1680  
Db 1621 TTCAAGAGGCTATGACTAGGTACTTGCCCCCCCCGGGAGACCCGCCCAACAGAAATAC 1680  
QY 1681 GACCTGAGTTGATPAACATATGCTCTCCCAATGTGTGCTGCGGACAGATGCAATTCGGC 1740  
Db 1681 GACCTGAGTTGATPAACATATGCTCTCCCAATGTGTGCTGCGGACAGATGCAATTCGGC 1740  
QY 1741 AAAAGGTAATACTACCTCAACCGGTGACCCG 1770  
Db 1741 AAAAGGTAATACTACCTCAACCGGTGACCCG 1770

RESULT 2  
A32184 1770 bp RNA linear PAT 10-DEC-1996  
LOCUS NAMBHV PT polypeptide fragment.  
DEFINITION NAMBHV PT polypeptide fragment.  
ACCESSION A32184  
VERSION A32184.1 GI:1926538  
KEYWORDS  
SOURCE  
ORGANISM Non-A, non-B hepatitis virus  
Non-A, non-B hepatitis virus  
Viruses; unclassified viruses.  
REFERENCE  
1 (bases 1 to 1770)  
AUTHORS  
JOURNAL  
FEATURES  
source  
Patent: FR 2655990-A 3 21-JUN-1991;  
Location/Qualifiers  
1..1770  
/organism="Non-A, non-B hepatitis virus"  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:12440"  
/codon\_start=1  
/product="partial NAMBHV-PT polypeptide"  
/protein\_id="CAA02088.1"  
/db\_xref="GI:1926539"  
/translation="QNDFPDADLIENLIMHEWKGDI TRVESKNKVILDSFPLRA

CDS



EEDEREVSPABILRSKFFPAMPAPARPDYNPILLESKAPDPVPPVHGGPLPT  
 KTRPIPPRRKTRVYLTESTVSSALAEIARFGSSGSDVSGTATAPDSDG  
 AGSDVSESSMPLREBPDPLSDSGSASOROKKATFDLQILDHYOVLEEMAK  
 EESKLPTINALSNILRRHNNVATRSRSATVSRKATFDLQILDHYOVLEEMAK  
 ASATVKAQLSVEEACKLTPHSAKREKYGAKRPAIIPDLGVRNLSKAIHISVMDLEDET  
 PITTTIAKNEVECVOPBERGRKPAIIPDLGVRNLSKAIHISVMDLEDET  
 YGROYSGORVEFLVNAKMSKTPMGVATRCFDSVTENDIRVESITVOCDDLAPE  
 ARQAIISLTERLYTIGPLTMSKNGKONGCYRRCASGVLTTSNGMTLTCYIKASAA  
 KLODCTMLVCGDLVITCEBAGTQEDNASLRFTETAMTRISAPGDPPEYLELIT  
 SCSNVSVAHDAAGKRVYILTRDP"

## ORIGIN

Query Match 100.0%; Score 1770; DB 6; Length 1770;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATGACTTCCAGACGCTGACCTCATCGAGGCCAACCCTGTGGCGGCATGAGATG 60  
 DB 1 CAAATGACTTCCAGACGCTGACCTCATCGAGGCCAACCCTGTGGCGGCATGAGATG 60

QY 61 GGGGGGACATTACCGCGCTGAGTCAAGAAACAAGTAACTCTTGACCTTTTGAC 120  
 DB 61 GGGGGGACATTACCGCGCTGAGTCAAGAAACAAGTAACTCTTGACCTTTTGAC 120

QY 121 CGGCTCGAGCGGAGAGATGAGCGGGAGTGTCCGTCGCGGCGAGATCTGCGGAAA 180  
 DB 121 CGGCTCGAGCGGAGAGATGAGCGGGAGTGTCCGTCGCGGCGAGATCTGCGGAAA 180

QY 181 TCCAGAAATTTCCACGACGATGCGCGCATGAGGCAAGCCGGAATTAACAACCTTCGCTG 240  
 DB 181 TCCAGAAATTTCCACGACGATGCGCGCATGAGGCAAGCCGGAATTAACAACCTTCGCTG 240

QY 241 CTGGAATCTCTGGAAGGCGCGGACACTAGTCTCTCAAGTGAATGAGTGGTCCCACTGCGCA 300  
 DB 241 CTGGAATCTCTGGAAGGCGCGGACACTAGTCTCTCAAGTGAATGAGTGGTCCCACTGCGCA 300

QY 301 CCTACTAAGACCCCTCTATACCACTTCAAGGAGGAGGACAGTGTCTTGAACGAA 360  
 DB 301 CCTACTAAGACCCCTCTATACCACTTCAAGGAGGAGGACAGTGTCTTGAACGAA 360

QY 361 TCCACGCTGTCTTCTGCGCTGCGGAGCTTGCACAAAGGCTTTTGATGCTCGGACG 420  
 DB 361 TCCACGCTGTCTTCTGCGCTGCGGAGCTTGCACAAAGGCTTTTGATGCTCGGACG 420

QY 421 TCGGCGGTGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAG 480  
 DB 421 TCGGCGGTGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAG 480

QY 481 GCGAGATCTGACGTTGAGTGTATCTCTCAATGCCCCCTTGAAGGGGAGCGGGGGAC 540  
 DB 481 GCGAGATCTGACGTTGAGTGTATCTCTCAATGCCCCCTTGAAGGGGAGCGGGGGAC 540

QY 541 CCCGATCTCAAGCAGGCTTGTGCTCAACGTAAGTGAAGGCGGCTGAGAGAGTGTGTC 600  
 DB 541 CCCGATCTCAAGCAGGCTTGTGCTCAACGTAAGTGAAGGCGGCTGAGAGAGTGTGTC 600

QY 601 TGTGCTCGATGTCTCAATGAGAGGCGGCTCTGATCAAGCCATCGCTGCGGAGGAA 660  
 DB 601 TGTGCTCGATGTCTCAATGAGAGGCGGCTCTGATCAAGCCATCGCTGCGGAGGAA 660

QY 661 AGCAAGTGGCCATCAACGCGTGAAGCACTTTTGTGCTGATCAACCAATGATGTCTAC 720  
 DB 661 AGCAAGTGGCCATCAACGCGTGAAGCACTTTTGTGCTGATCAACCAATGATGTCTAC 720

QY 721 GCTACACATCCGAGCGGCAAGCGGCAAGGAGTCACTTTTGAACAGCTGCA 780  
 DB 721 GCTACACATCCGAGCGGCAAGCGGCAAGGAGTCACTTTTGAACAGCTGCA 780

QY 781 ATCTTGAAGATCACTACAGAGAGTGTCTCAAGAGATGAAGGAGGAGGCTTCAAGCTT 840  
 DB 781 ATCTTGAAGATCACTACAGAGAGTGTCTCAAGAGATGAAGGAGGAGGCTTCAAGCTT 840

QY 841 AAGGCTAAGCTTCTATCAATGAGAGAGGCTGCAAGCTGACGCCCCCACTTGGGCCAAA 900

DB 841 AAGGCTAAGCTTCTATCAATGAGAGAGGCTGCAAGCTGACGCCCCCACTTGGGCCAAA 900

QY 901 TCTAAATTTGGCTATGAGGCAAGAGAGTCCGGAACCTTATCCAGAGGCAATTAACAC 960  
 DB 901 TCTAAATTTGGCTATGAGGCAAGAGAGTCCGGAACCTTATCCAGAGGCAATTAACAC 960

QY 961 ATCCGCTCGTGTGAGAGAGCTTGTGGAAGACCTGAAGAACCAATTTGACACACATC 1020  
 DB 961 ATCCGCTCGTGTGAGAGAGCTTGTGGAAGACCTGAAGAACCAATTTGACACACATC 1020

QY 1021 ATGGCAAAAATGAGGTTTTCTGCTCAACAGAGAGAGAGGCGGCAAGCCAGCTGCG 1080  
 DB 1021 ATGGCAAAAATGAGGTTTTCTGCTCAACAGAGAGAGAGGCGGCAAGCCAGCTGCG 1080

QY 1081 CTATGCTGTCTCCAGACTTTGGGGGTCGCTGTGTGTGGAAGAAATGAGCCCTTATGACGTG 1140  
 DB 1081 CTATGCTGTCTCCAGACTTTGGGGGTCGCTGTGTGTGGAAGAAATGAGCCCTTATGACGTG 1140

QY 1141 GTCTCAACCTCCCTGAGGCTGATGAGGCTCCTGTAAGGATTCAGATTCCTCTGGA 1200  
 DB 1141 GTCTCAACCTCCCTGAGGCTGATGAGGCTCCTGTAAGGATTCAGATTCCTCTGGA 1200

QY 1201 CAGCGGCTGAGTCTCTGCTGAACGCTGGAATCAAAAGAGACCCCTATGAGCTTTGGA 1260  
 DB 1201 CAGCGGCTGAGTCTCTGCTGAACGCTGGAATCAAAAGAGACCCCTATGAGCTTTGGA 1260

QY 1261 TATGACACCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTGTAGAGAGTCA 1320  
 DB 1261 TATGACACCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTGTAGAGAGTCA 1320

QY 1321 ATTATCAATGTTGATCTTGGCGGCGGAGCGGCAAGGCTTAAGTGTGCTCAAGAG 1380  
 DB 1321 ATTATCAATGTTGATCTTGGCGGCGGAGCGGCAAGGCTTAAGTGTGCTCAAGAG 1380

QY 1381 CGGCTTTATATCGGGGCTCCCTGACTAATTTCAAAAGGCGAGAACTGCGGCTTATCGCGG 1440  
 DB 1381 CGGCTTTATATCGGGGCTCCCTGACTAATTTCAAAAGGCGAGAACTGCGGCTTATCGCGG 1440

QY 1441 TGCAGCGGAGCGGCTGTGACGACTGAGTGTGTAATACCTCATGATTTACTTTGAAG 1500  
 DB 1441 TGCAGCGGAGCGGCTGTGACGACTGAGTGTGTAATACCTCATGATTTACTTTGAAG 1500

QY 1501 GCTCTGACGCTGTGTGAGCTGCAAGGCTTCAAGACTGACAGTGTGCTGTGCGGAGAG 1560  
 DB 1501 GCTCTGACGCTGTGTGAGCTGCAAGGCTTCAAGACTGACAGTGTGCTGTGCGGAGAG 1560

QY 1561 GGCCTGTGCTTATCTGTGAGAGCGGAGAACCCAGAGAGAGCGGCGAGGCTTACAGAGTC 1620  
 DB 1561 GGCCTGTGCTTATCTGTGAGAGCGGAGAACCCAGAGAGAGCGGCGAGGCTTACAGAGTC 1620

QY 1621 TTCACGAGAGCTATGACTGAGTACTTGTGCCCCCGGAGAGCCGCGCCAAACCAAGATAC 1680  
 DB 1621 TTCACGAGAGCTATGACTGAGTACTTGTGCCCCCGGAGAGCCGCGCCAAACCAAGATAC 1680

QY 1681 GACCTGAGATTTGATTAATCAATGCTCTCAATGTGTGCTGTGCGGCAAGATGATCTGTGC 1740  
 DB 1681 GACCTGAGATTTGATTAATCAATGCTCTCAATGTGTGCTGTGCGGCAAGATGATCTGTGC 1740

QY 1741 AAAAGGATTAATCACTCAACCGGTGACCGG 1770  
 DB 1741 AAAAGGATTAATCACTCAACCGGTGACCGG 1770

RESULT 3  
 AR144032  
 LOCUS AR144032 1770 bp DNA linear PAT 08-AUG-2001  
 DEFINITION Sequence 3 from patent US 6210675.  
 ACCESSION AR144032  
 VERSION AR144032.1 GI:15105899  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.



other sequences; artificial sequences.  
1 (bases 1 to 3750)

REFERENCE  
AUTHORS Post-translational non-A non-B hepatitis viral polypeptides  
TITLE Patent: GB 2239245-A 22 JUN-1991;  
JOURNAL Location/Qualifiers  
FEATURES  
Source 1. 3750

1. 3750  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="contig formed by cDNA clones from 3' end"  
<1..3750  
/codon\_start=1  
/rname="table=1"  
/product="PT-NANBH polypeptide"  
/protein\_id="CA01932.1"  
/db\_xref="GI:1248641"

CDS

PSMDOMKCLIRLKPRLRGPTPLLYRLGAVQNEVTLHPITKTMCMGADLEVVST  
WLVGVVLAALAYCLTGSVVIVGRILISGRPAIYDPDEVLYQEPDEMECSHLY  
IEGMOLAEQFKOKALGLQATKQABAAPVESKRALLETWMAHNNFISGTYL  
AGLSTLPGNPALASLMAFTASVTPLTOSTLLNLIGLWVAQLAPPSAASFVAG  
IAGAAGSGLGKVLVDIILAGVAGVAGALVAFKMGSEMPSTEDLVNLLPALISFGA  
LVYVCAALIRRHVGPBEGA VQMMNRLLAFASRGNNVSPTHVPESDAARVOTLS  
DLITTLKRLHQMINDCSTPGSGMLRDVMDICTVLADFTWLQSLPLPGVP  
FSCQKGYKWRGDSIMOTTSCGAQITGHVNGSMRIYGPCTCSMMHRTPIINAY  
TTGCTPSAPNTSRALMRYAAEYEVTRKVDVDFHVTSTTDNVACPCQVPAPEFT  
EVDGRLHRYAPACKPLREVEYQVGLVYVSQLPEPEPDAVLVLSMLDPBHI  
TAEAKRLRAGSPSLASSASOLASPSKATYITONFPDPLLEMLMHEMG  
DIATRESENKVVILDSFDPRAEDEREVSYPALIRKSKPPAMPAMRPVNP  
LESWKAADVVPVHGCPLEPTKPIIPERRKRVTLRESYTSALALATAPGSS  
BRSAVDSGRITAPDPDPSDGGAGSVESISNRPLEGEGBDLSDSKSTYSBAG  
EDVACSSMTWTGALITPCAABESKLPINALSNSLLRHNNVATLTSBASQKQV  
TEPDRLIDLDYQDVLKEMKAKATVAKKLVEBAKCLTPPSASKSGYAKDVRN  
LSKAIINHRSVWEDLEDETETPITDTIAKNEVFCVQERGSKKPAIIVFDLCVR  
VCEKMLADYVSTLPQAMGSSYGFQYSPQORVEPLVNMVSKTGMGFAYDRGDS  
TYTENDIRVESIYOCCLDAPAROAIRSLTERLYIGPLTSTKGNCGRRRBAAGV  
LTTSCNTLTCYIKASAPRAALQDCTMIVCGDGLVIVICESAGTEBDASLRFVTEA  
MTRISAPGDPPEYDELITSSNSVSAHDASGRVYLTDP

## ORIGIN

Query Match 99.64; Score 1763.6; DB 6; Length 3750;  
Best Local Similarity 99.64; Pred. No. 0;  
Matches 1766; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 CAAAATGACTTCCCAAGCTGACCTCATGAGAGCCAACTCTGTGGCGGCAATGAGATG 60  
DB 1981 CAAAATGACTTCCCAAGCTGACCTCATGAGAGCCAACTCTGTGGCGGCAATGAGATG 2040  
QY 61 GGGGGGGAATTCCCGCGGTGAGTCAAGAGAAAGAGTAATCTCTGGAATCTTTTGAC 120  
DB 2041 GGGGGGGAATTCCCGCGGTGAGTCAAGAGAAAGAGTAATCTCTGGAATCTTTTGAC 2100  
QY 121 CGGCTCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCGCGGCGGAGATCTGCGGAAA 180  
DB 2101 CGGCTCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCGCGGCGGAGATCTGCGGAAA 2160  
QY 181 TCCAGAAATTTCCACAGGAGATGCGCGCAATGAGCAAGCGCGGATTAACAACCTTCGCTG 240  
DB 2161 TCCAGAAATTTCCACAGGAGATGCGCGCAATGAGCAAGCGCGGATTAACAACCTTCGCTG 2220  
QY 241 CTGGAATCTCTGGAAGGCGCGGAGTACGTCTTCAGTGTGTAATGAGTGGTCCCACTGGCA 300  
DB 2221 CTGGAATCTCTGGAAGGCGCGGAGTACGTCTTCAGTGTGTAATGAGTGGTCCCACTGGCA 2280  
QY 301 CCTACTAAGACCCCTCTCTATACCACTCCAGGAGAAAGAGACAGTGTGTTGACAGAA 360  
DB 2281 CCTACTAAGACCCCTCTCTATACCACTCCAGGAGAAAGAGACAGTGTGTTGTTCTGACAGAA 2340  
QY 361 TCCACGCTGTCTTCTGCGCTGAGGAGCTTGCAAAAGGCTTTTGGTAGCTCCGAGCG 420  
DB 2341 TCCACGCTGTCTTCTGCGCTGAGGAGCTTGCAAAAGGCTTTTGGTAGCTCCGAGCG 2400

QY 421 TGGCCGTGAGACGAGCGGCAAGGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGCA 480  
DB 2401 TGGCCGTGAGACGAGCGGCAAGGCAACGCGCCCTCTGACCAATCTCCGACGAGCGGCA 2460  
QY 481 GCAGATCTGACGTTGAGTCTGATATCTCTCATGCCCCCTTTGAGGGGAGAGCGGAGGAC 540  
DB 2461 GCAGATCTGACGTTGAGTCTGATATCTCTCATGCCCCCTTTGAGGGGAGAGCGGAGGAC 2520  
QY 541 CCCGATCTGACGAGCGGAGTCTGTTGTTACCGGTGAGTGAAGAGCGGTGAGGACCTGCTC 600  
DB 2521 CCCGATCTGACGAGCGGAGTCTGTTGTTACCGGTGAGTGAAGAGCGGTGAGGACCTGCTC 2580  
QY 601 TGCTGCTGATGTCCTACACATGAGCAGGAGCTGTATACCGCATAGCGCTGCGAGGAA 660  
DB 2581 TGCTGCTGATGTCCTACACATGAGCAGGAGCTGTATACCGCATAGCGCTGCGAGGAA 2640  
QY 661 AGCAAGCTCCCATCAACGCGTTGAGCAACTCTTGTGCTGCTGACCAACAATGCTCTAC 720  
DB 2641 AGCAAGCTCCCATCAACGCGTTGAGCAACTCTTGTGCTGCTGACCAACAATGCTCTAC 2700  
QY 721 GCTACCAATCCCGAGCGGCAAGCCAGGCGGAGAGAGAGTCACTTTGACAGACTGCA 780  
DB 2701 GCTACCAATCCCGAGCGGCAAGCCAGGCGGAGAGAGAGTCACTTTGACAGACTGCA 2760  
QY 781 ATCTGGAAGATCACTACAGGACGTCCTCAAGAGATGAGGCGAAGGCGTCCACAGTT 840  
DB 2761 ATCTGGAAGATCACTACAGGACGTCCTCAAGAGATGAGGCGAAGGCGTCCACAGTT 2820  
QY 841 AAGGCTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACCGCCCATTCGAGCA 900  
DB 2821 AAGGCTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACCGCCCATTCGAGCA 2880  
QY 901 TCTAAATTTGGCTATGAGGCAAGAGCGTCCGAACTTATCCAGCAAGCCATTAAACAC 960  
DB 2881 TCTAAATTTGGCTATGAGGCAAGAGCGTCCGAACTTATCCAGCAAGCCATTAAACAC 2940  
QY 961 ATCCGCTCGGTGAGGAGCACTTGTGGAAGACACTGAACACCAATTGACACCAATC 1020  
DB 2941 ATCCGCTCGGTGAGGAGCACTTGTGGAAGACACTGAACACCAATTGACACCAATC 3000  
QY 1021 ATGCAAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGAGCGGCAAGCAGCTCCG 1080  
DB 3001 ATGCAAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGAGCGGCAAGCAGCTCCG 3060  
QY 1081 CTATGCTGTTCCAGACTTGGGAGTCCGTGTGTGAGAGAAATGAGCCCTTATGACGTG 1140  
DB 3121 CTATGCTGTTCCAGACTTGGGAGTCCGTGTGTGAGAGAAATGAGCCCTTATGACGTG 3180  
QY 1141 GTCTCAACCTCCAGAGCTGAGATGAGGCTCTGTAAGGATTCAGATTCCTCTGGA 1200  
DB 3121 GTCTCAACCTCCAGAGCTGAGATGAGGCTCTGTAAGGATTCAGATTCCTCTGGA 3180  
QY 1201 CAGCGGCTGAGATTCTGTGTAAGCGCTGGAATCAAGAAAGCCCTATGAGCTTTGCA 1260  
DB 3181 CAGCGGCTGAGATTCTGTGTAAGCGCTGGAATCAAGAAAGCCCTATGAGCTTTGCA 3240  
QY 1261 TATGACACCCGCTGTTTGAATCAACAGTCACTGAGATGACATTCGTGTAGAGATCA 1320  
DB 3241 TATGACACCCGCTGTTTGAATCAACAGTCACTGAGATGACATTCGTGTAGAGATCA 3300  
QY 1321 ATTATGATGTTTGAATCTTGGCCCCCGGAGGCAAGCCATTAAGTGTGCTCAAGAG 1380  
DB 3301 ATTATGATGTTTGAATCTTGGCCCCCGGAGGCAAGCCATTAAGTGTGCTCAAGAG 3360  
QY 1381 CGGCTTATATGAGGAGTCCCTGACTAATTAATAAAGGCAAGCTGAGGCTATGCGCGG 1440  
DB 3361 CGGCTTATATGAGGAGTCCCTGACTAATTAATAAAGGCAAGCTGAGGCTATGCGCGG 3420  
QY 1441 TGCAGGAGAGCGGCTGTGACAGATGAGTGCAGTATACCTTCAATGTTACTTTAAG 1500  
DB 3421 TGCAGGAGAGCGGCTGTGACAGATGAGTGCAGTATACCTTCAATGTTACTTTAAG 3480  
QY 1501 GCCTGTGACGCTGTGAGACTGCAGAAAGCTCCAGAGCTGACAGATGCTGCTGTGCGAGAC 1560

Db 3481 GCGCTGACGCGCTGAGCTGCAAGCTGCAAGACTGCTGTGGCGAGAC 3540  
Qy 1561 GGGCTGTGTTATCTGTGAGAGCGGGAAACCCAGAGAGCGCGCGCTTACAGAGTC 1620  
Db 3541 GGGCTGTGTTATCTGTGAGAGCGGGAAACCCAGAGAGCGCGCGCTTACAGAGTC 3600  
Qy 1621 TTCACGAGAGCTATGACTGACTGTGCCCCCGCGGAGACCGCGCCAAACAGAAATAC 1680  
Db 3601 TTCAGGAGAGCTATGACTGACTGTGCCCCCGCGGAGACCGCGCCAAACAGAAATAC 3660  
Qy 1681 GACCTGAGAGTGAATCAATGATGCTCTCCAAATGTGTGCGTGGCGACGATGATCTGGC 1740  
Db 3661 GACCTGAGAGTGAATCAATGATGCTCTCCAAATGTGTGCGTGGCGACGATGATCTGGC 3720  
Qy 1741 AAAAGGATATGACTGACTGACCGGTGACCGG 1770  
Db 3721 AAAAGGATATGACTGACTGACCGGTGACCGG 3750

RESULT 5  
A32203  
LOCUS A32203 3750 bp DNA linear PAT 10-DEC-1996  
DEFINITION NANBH PT polyprotein fragment.  
ACCESSION A32203  
VERSION A32203.1 GI:1926564  
KEYWORDS  
SOURCE Non-A, non-B hepatitis virus  
ORGANISM Non-A, non-B hepatitis virus  
REFERENCE Viruses; unclassified viruses.  
1 (bases 1 to 3750)  
AUTHORS Patent: FR 2655990-A 22.21-JUN-1991;  
JOURNAL location/Qualifiers  
FEATURES  
source 1..3750  
/organism="Non-A, non-B hepatitis virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:12440"  
<1..>3750  
/codon\_start=1  
/product="partial NANBH-PT polyprotein"  
/protein\_id="CAA02095.1"  
/db\_xref="GI:1926565"

CDS

/translation="MEGVFTGLTHVDAHFLSQTQAGDNFPLYVAQATVCARAQAP  
PSMDMMKCLIRLKLTLRGPTPLRYLGAVONEVTLTHPTITFIACMSADLEVTST  
WLVGVGVAALAAVCLTGSVYIVGRILLGSRPAIVREVIVYQSDMEASRLPY  
IEQGMQAEQKQXALGLIQTATKQAEAAVYESKMRLETFAKHMNFTSGIOYL  
AGLSTDPGNPAIASLMAFTASVTSPLTQSTLLNLILGAVPNLSLPASAFVAG  
IAGAAGVSGIGAKIVLDILAGAGVAGALVAFKVMGSMPTEDLVNLPALISGA  
LVGVVCAAILRRHVPGEAGVQMMNRLLAFASRGNHSPHYVESDAAARVQILS  
DLITLQRLKHLHINEDSTPGSGSMRLRDVMDICTVLADKTMLOSGLAPRLGVP  
FFSCRGYGVWVRGIMQTTGSCAOLITGVKNGSMITGVKTCSSNMHGFPIINAY  
TTGPTCPSPAPVSRALNRVAEEVEYTRVGDHYVTSMITDNKCPQVPAPFET  
EVDGRLHRYAPACKPLREEVTFQVGNQYLVGSQLPEPPDVAVALTSMITDESHI  
TAETAKRLRARGSPSLSSASQLSASBSKATYITQNDFPADILRNLRLRMHMG  
DITRVESKRVILDSFPLRAEEDEREVSVAELIKSKKPPAMPAMPARDPYPL  
LESMAADVPVPHVGCPLPTKPTPIPPRRKRVVLTESVSLALBELATKARGSS  
BBSAIVDSGATAPADPOPSDNDGAGSDVSYSMPLEGRPDPDLSGMSVTSBAG  
EDVVCSSMYTGTALITPCAEBSKLPINALSNSLRLHNNVIVITRSASQKQV  
TSDRLQILDDHYODVLKEMKAKASTVKALSVESACKLTPHSAKSKFEGAGAVRN  
LSKAIINHRSVMEIDLBDTEPITDTTMAKNEVCVOPBERGKRPALIIYFDPKVR  
VCEKALYDVVSTLFOAVGSSYGFQYSGQVEFLVNAKSKTGMGFAVDRCFDS  
TYTENDIRVESITVCCDLAPBAROIRISLTERLYIGGLTMSKONGSPYRCRASGV  
LTTSGNLTLCYLAACRAKLODKTMVCGDLIVYI CESAGTQEDBAASLRVTTAA  
MTRYSAPDEDPQPEYDIELITSCSSNSVAAHDAGKRYIYITRDP"

ORIGIN

Query Match 99.6%; Score 1763.6; DB 6; Length 3750;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1766; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAAAATGACTTCCAGAGCGCTGACTGACGCGCAACCTCTGTGGCGGATGAGATG 60  
|||||  
Db 1 CAAAATGACTTCCAGAGCGCTGACTGACGCGCAACCTCTGTGGCGGATGAGATG 60

Db 1981 CAAAATGACTTCCAGAGCGCTGACTGACGCGCAACCTCTGTGGCGGATGAGATG 2040  
Qy 61 GCGCGGGACATTACCCCGGTGAGAGTCAAGAAACAAGTAATCTGAGACCTTTGAC 120  
Db 2041 GCGCGGGACATTACCCCGGTGAGAGTCAAGAAACAAGTAATCTGAGACCTTTGAC 2100  
Qy 121 CCGCTCCGAGGAGAGAGATGAGCGGAGATGTCTCCGCGCGAGATCTCGCGAAA 180  
Db 2101 CCGCTCCGAGGAGAGAGATGAGCGGAGATGTCTCCGCGCGAGATCTCGCGAAA 2160  
Qy 181 TCAGAAATTCACACAGCATGCGCGATGAGGACCGCGGATTAACAACCTCGCTG 240  
Db 2161 TCAGAAATTCACACAGCATGCGCGATGAGGACCGCGGATTAACAACCTCGCTG 2220  
Qy 241 CTGAGATCTGGAAGGCGCGGAGCTAGAGTCCCTCAGTGGTACATGGGTGCCATGCA 300  
Db 2221 CTGAGATCTGGAAGGCGCGGAGCTAGAGTCCCTCAGTGGTACATGGGTGCCATGCA 2280  
Qy 301 CCTACTAAGACCCCTCTATACACCTTCACAGAGAAAGAGACAGTTGTTCTGACAGAA 360  
Db 2281 CCTACTAAGACCCCTCTATACACCTTCACAGAGAAAGAGACAGTTGTTCTGACAGAA 2340  
Qy 361 TCACCGTGTCTTCTGCGCTGAGCGAGCTTGCACAAAGCTTTTGTAGCTCCGAGACG 420  
Db 2341 TCACCGTGTCTTCTGCGCTGAGCGAGCTTGCACAAAGCTTTTGTAGCTCCGAGACG 2400  
Qy 421 TCGCGCGTTCAGACGAGGAGCGCAACCGCGCTCTGACCAATCTCCGAGAGCGGGA 480  
Db 2401 TCGCGCGTTCAGACGAGGAGCGCAACCGCGCTCTGACCAATCTCCGAGAGCGGGA 2460  
Qy 481 GCAGATCTGACGTTGATGCTATATCTCCATGCCCCCTTGAGGAGGAGCCGAGGAC 540  
Db 2461 GCAGATCTGACGTTGATGCTATATCTCCATGCCCCCTTGAGGAGGAGCCGAGGAC 2520  
Qy 541 CCCGATCTCAGCGAGCGGCTTTGGTCTTACCGTGAAGAGAGCGCGTGAAGAGTCTG 600  
Db 2521 CCCGATCTCAGCGAGCGGCTTTGGTCTTACCGTGAAGAGAGCGCGTGAAGAGTCTG 2580  
Qy 601 TGCGTCTGATGCTCTACATGAGCAGGCGCTGATCAAGCCATGCGTGGAGAGAA 660  
Db 2581 TGCGTCTGATGCTCTACATGAGCAGGCGCTGATCAAGCCATGCGTGGAGAGAA 2640  
Qy 661 AGCAAGCTGCCATCAACGCGTTGAGCAACTTTTGTGCTGCTCAACAACATGCTTAC 720  
Db 2641 AGCAAGCTGCCATCAACGCGTTGAGCAACTTTTGTGCTGCTCAACAACATGCTTAC 2700  
Qy 721 GCTACCATATCCGAGCGGACGACGCGAGCGAGAAAGGTCACTTTGACAGACTGCA 780  
Db 2701 GCTACCATATCCGAGCGGACGACGCGAGCGAGAAAGGTCACTTTGACAGACTGCA 2760  
Qy 781 ATCCGAGCGATCACTACAGAGAGTGTCAAGAGATGAAGGAGGCGTCCACAGTT 840  
Db 2761 ATCCGAGCGATCACTACAGAGAGTGTCAAGAGATGAAGGAGGCGTCCACAGTT 2820  
Qy 841 AAGGCTAAGCTTCTATCAGTGAAGAAAGCTGCAAGCTGACGCCCCCACTTCCGCCAAA 900  
Db 2821 AAGGCTAAGCTTCTATCAGTGAAGAAAGCTGCAAGCTGACGCCCCCACTTCCGCCAAA 2880  
Qy 901 TCTAATTTGGCTATGAGGAGCAAGAGAGCTCCGAAACCTATCAGACAAAGCCATTACAC 960  
Db 2881 TCTAATTTGGCTATGAGGAGCAAGAGAGCTCCGAAACCTATCAGACAAAGCCATTACAC 2940  
Qy 961 ATCCGCTCCGAGTGGAGAGCTTGTGGAACAACCTGAACAACAATTGACACACATC 1020  
Db 2941 ATCCGCTCCGAGTGGAGAGCTTGTGGAACAACCTGAACAACAATTGACACACATC 3000  
Qy 1021 ATGCAAAAAATGAGGTTTCTGCGTCCACACAGAGAGAGGCGCGCAACCGAGCTGCG 1080  
Db 3001 ATGCAAAAAATGAGGTTTCTGCGTCCACACAGAGAGAGGCGCGCAACCGAGCTGCG 3060  
Qy 1081 CTTATCGTTCCTCAGACTTGGGAGTCCGTGTGTGCGAGAAATAGCCCTTATGACGTG 1140  
Db 3061 CTTATCGTTCCTCAGACTTGGGAGTCCGTGTGTGCGAGAAATAGCCCTTATGACGTG 3120



```

Qy 1141 GTCCTCACCCTCCCTCAGGCTGTGATGGGCTCTCTGTAAGGATTTCAGTATTCCTCTGGA 1200
    |||||
Db 3121 GTCCTCACCCTCCCTCAGGCTGTGATGGGCTCTCTGTAAGGATTTCAGTATTCCTCTGGA 3180
Qy 1201 CAGCGGCTGAGATTCTGTGTAAACGCTTGAAATCAAAGAAGACCCCTATGCGCTTTGGA 1260
    |||||
Db 3181 CAGCGGCTGAGATTCTGTGTAAACGCTTGAAATCAAAGAAGACCCCTATGCGCTTTGGA 3240
Qy 1261 TATGACACCCGCTGTTTGAAGTCAACAGTCACTGAGGAATGATCCGTGTAGAGAGTCA 1320
    |||||
Db 3241 TATGACACCCGCTGTTTGAAGTCAACAGTCACTGAGGAATGATCCGTGTAGAGAGTCA 3300
Qy 1321 ATTATCAATGTTGTGACTTTGGCCCCCGAAGCAGACAGCCATTAAGTGTGCTCAGAG 1380
    |||||
Db 3301 ATTATCAATGTTGTGACTTTGGCCCCCGAAGCAGACAGCCATTAAGTGTGCTCAGAG 3360
Qy 1381 CGGCTTATATGCGGGGTCCCTGACTAATTCAAAGGCGAGACTGCGGCTATGCGCG 1440
    |||||
Db 3361 CGGCTTATATGCGGGGTCCCTGACTAATTCAAAGGCGAGACTGCGGCTATGCGCG 3420
Qy 1441 TGCCTGCGAGGCGGCGGTGCTGAAGACTAGCTGGGGAATACCTCAATGTTACTTTGAAG 1500
    |||||
Db 3421 TGCCTGCGAGGCGGCGGTGCTGAAGACTAGCTGGGGAATACCTCAATGTTACTTTGAAG 3480
Qy 1501 GCCTCTGACGCTGTGAGCTGCAAGCTCAGAGCTGACAGATGCTGCTGTGCGGAGAC 1560
    |||||
Db 3481 GCCTCTGACGCTGTGAGCTGCAAGCTCAGAGCTGACAGATGCTGCTGTGCGGAGAC 3540
Qy 1561 GGCCTTGTGCTTATCTGTGAGAGCGCGGGAACCCAGAGAGACGCGCGGAGCTTACGAGTC 1620
    |||||
Db 3541 GGCCTTGTGCTTATCTGTGAGAGCGCGGGAACCCAGAGAGACGCGCGGAGCTTACGAGTC 3600
Qy 1621 TTCAGGAGGCTATGACTAGTACTGCTGCCCCCGCGGAGACCCGCCCAACAGAAATAC 1680
    |||||
Db 3601 TTCAGGAGGCTATGACTAGTACTGCTGCCCCCGCGGAGACCCGCCCAACAGAAATAC 3660
Qy 1681 GACCTGAGATTGATTAACATCATGCTCTTCCAAATGTCGTCGCGCAGCATGATCTGGC 1740
    |||||
Db 3661 GACCTGAGATTGATTAACATCATGCTCTTCCAAATGTCGTCGCGCAGCATGATCTGGC 3720
Qy 1741 AAAAGGTTACTACTCTCAACCCGTGACCCG 1770
    |||||
Db 3721 AAAAGGTTACTACTCTCAACCCGTGACCCG 3750

```

```

RESULT 7
AF313916 9359 bp mRNA linear VRL 01-JAN-2002
LOCUS Hepatitis C virus polyprotein mRNA, partial cds.
DEFINITION AF313916
ACCESSION AF313916
VERSION AF313916.1 GI:18027684
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
          Hepacivirus.
REFERENCE 1 (bases 1 to 9359)
AUTHORS Fanning, L.J., Itakura, J., Nagayama, K. and Enomoto, N.
TITLE Characteristics of Hepatitis C viral genome associated with disease
        progression in a homogeneous patient population
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 9359)
AUTHORS Fanning, L.J., Itakura, J., Nagayama, K. and Enomoto, N.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2000) Medicine, National University of Ireland,
        Cork, Hepatitis C Unit, Clinical Sciences Building, Cork University
        Hospital, Cork, Ireland
        location/Qualifiers
FEATURES
source 1..9359
        /organism="Hepatitis C virus"
        /mol_type="mRNA"
        /db_xref="taxon:11103"

```

5'UTR  
CDS

```

/note="Isolated from source plasma identified as the
single source of the iatrogenic infection of Rhesus
negative women in Ireland from May 1977-November 1978;
exposure was through HCV 1b contaminated anti-D
immunoglobulin"
1..329
330..9359
/codon_start=1
/product="polyprotein"
/protein_id="AAL55821.1"
/db_xref="GI:18027685"
/translation="MSTNPKPQRTKNTNRPPDVKFPGGGQIVGVYLLIPRGPRL
GVARTKTSERSQPRGRQPIPKAREGRVMAQPGVWPVLENGMGMAWMLSPRG
SRPSPGPDPRRRRNILKVIDTLCTGADLMGIPLVAGALGGAALAGVLEED
GVNATGNLPGCSRSLFILALISGLTIPASILBVRNYSVAVHNDSCNASIVLEAD
MIMHPGCVPCVRENNSRCVNLPLPLTANNSILPTTIRRHDLIVGTAPCSAMY
VGDLCGFTVLVSQLFPSRRHETVQDNCISYIFGHISGRHMADMMNNSPTLAVY
SOLIRIPQAVVDVAGAHVGLAGLAYSMAGNNAKVLIVNLLFAGVDEGTYYGAG
AHTTGLATLFSFSPSONIDINTGSMHINRLKNCNDSLNTGFLAALFYARFNAS
GCPERMASCRPIDKFAQMGPIAYAKPLSDQKCYMHYAPQCGIVASQVCPVYC
FTSPVVVGTTRDFGVPYTSWGENETDVLNTRPRPGNMFGCTMMNGTFTKCGG
PCDIDGVGNDTLCPDPCFKHPEATYTKGSGPWLTPRCMDYPRYLRHMPCTVNF
TIFKRYTVGVGHERLNAACWMTGRCDDLEDPRSLSPLLSITTEKQVLPCSFTHL
PALSTGLIHLHRLNIVDQYLYGISAASVTAIKREYVLLFLFLADAVCAIWMMLL
IVQEAALENLVNMAASVAGTNGILSEFLVEFCAMYIKRLVGAAYAFPGVPLLL
LILALPRRAYAMDREMAASCGAVFIGLALITLSPHYKFLARLIMVLOYETRAEAL
LOVNIPLPLNVRGSDAILILMCAIHPELIFATIKLILALIGLPLMLVAGTRVFPVR
AHGILRAVCVIVREAGHYVOMALMRLALNGTVYVHPLRMAHGRRLDALVAVR
PVFSDMETKILITWGADTAACGDIILGLPYSARGREILLAGPADSLSCQRLAIPIT
AVAOQTRGLGCLIVTSLGRDKNVEBQVAVSTATOSPLATCVNGCVYVHAGATK
TLAQKGPITOMYTNVODULVGMPLAPGARSLTPCTGSSDLVLTTHADVTPVRRG
DSRSLSPRPVSYLKSSSGPLCPSPAGANGIIRAVCTGVAKAVDPVVESEMETT
MRSPVFTDNSSPPAVPOTFOYALHAPITGSGSKSPVAAVAAQGVKVLVNLPSVATL
GFGAMKKAHVDPNIRTVGRITTTGAPIYTSYGRKLADGCGGAGVDIITCDECHS
TDSITLIGITVLDQARTAGRLVLTATPDPGSVTPHNPTEVLSNIGETIPYTK
AIPETIKGGHLLFCHSRKCCDELAFLSGLNAAVAYKRGDVSVIPSGDVAVVA
TDALMTGFTGDFSDVIDCNCVTQTVFSLDPTPIETTVTPDQAVSRQSRGRTHV
RTGAYRRTVPEERPSGMFSDSYLCECYDAGCAEMVLAETSVYLRAYLTPGLPVC
DHLFEMGVFTGLTHIDAHPLSOTKRGDNPPYVLAOATVCARAOAPPPMDWMKC
LIRKPLTHGPTLLYTHIAQVONENVITLTHPTIKTMMQADLVEASTVTLVGSGLA
LILAYCITTSQSVIVGRIILSGKAPVDPREVLYOQDENBECKTHLPYIEQQLAB
QFKQKALGLLQATKQAEAAAPVESKRALSESTWAKHMMVFISGIVLADSLPLGN
PALASLMAFTASITSPLTQHTLLFNILAGVAAQALAPSAAPVAGAGAGAVGSI
GLAGLVLDIILAGYAGVAGALVAFKNSGEMSTEDVNLPLILISGALVGVCAA
ILRHVPGREGAVQMMNRLIAFASRGHVSPTHYVPSDAAARVTLISLITLQKX
RIHOWINEDCSTPSSGMLDVMWMTCTVLTDPFTWLOSLPLPLPVPVPCSGRGX
GVWFGDGIQOTTCCGQOITGHVKNKSMRIVGPTCSTNTHRTPTPIINAYTTPGCTPSP
APNYSRALMRTAEVLEVTQGDPHVYTGCTTDNVCCPCQVPAPEPTEDEGRRLR
YAPPCKPLREBEVTVQGLNOYLVGSQLPCPEPDDVAVLTSMLTDPESHITAEAKRL
ARGSPSLASSASQSLAPSLKATCTRHSDPDADILEANLIMQEMGINTRVESN
KVVLIDFEPFLABEDREVSVPAILIRSRKFPFAMPVIMRPPYNPLVSWMDPY
VPPVYHGGCLPPTKAPPIPPRRKRTVLTSTVYSALSLAELATKTPSSSSASDST
ATTAPDPSDDGAGSDVSESNRPVLEGEGRDPLSDGSRSTVSEBESDVYCCNS
YTWTGALITPCSAESKLPINALNSILRHHNVYATTSRSASQKRVFTRQLQVLD
DHYRVDLKEMKAKASTVAKLLPVEBAKCTKPPISAKSFGYDADRVNRSKXVNH
LSWVKDLLEDTEPIIDTTIMAKNEVFCQEGEKSPKAPRILVFEDLGRVCEKALVD
VSTLTPQAVNGSSYGFQYSPGQRYEFLVNMKKKNGMGAFTDRCDSTVTESDIIV
ESTVYOCCLAPPEARQIRSLTRELTVIGSLTKNGKNGCCRRRACGATVTSQNTL
TCYIKASACRAKLDQCTMLVCGDVLVYICESAGTIEDASLRFVTEAMTRYSAPRG
DPPPEPVDLELITSSCNVSVADHDSKRYVYLRDPTPLARAAMTARTHTVIAORLH
GNIMVAPPTLMARMIIMTHFFSILALQOEKLDLCOITYACVSIIBLDPOLIORLH
GLSAFSLHSYSPGEINVAACRLGLGVPLKEMHRRARSYAKLLISGGGRANCGKYL
ENMVRTRTKLKTPIPASQDLSWVFAVSGGDIYHSLBRARRRWMLCLLLISGV
GIYLLPRR"

```

## ORIGIN

```

Query Match 89.3%, Score 1580.6; DB 14; Length 9359;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
Qy 5 ATGACTTCCAGACGCTGACTCTATCGAGCGCAACTCTCTGTGCGCGCATAGATGGCGG 64
Db 6985 ATGACTCCCGCAGATGCTGACTCATCGAGCGCAATCTCTGTGCGCGCAGAGATGGCGG 7044

```



65 GGGACATTAACCCGCGGAGTCAAGAAAGTAGTAATCTGACTCTTTGACCCCG 124  
7045 GGAACATCACCCGCGGAGTCAAGAAAGTAGTAATCTTTGACCCCG 7104  
125 TCCGACGCGAGAGATGACCGGAAAGTCCGTCGCGGAGATCTTCCGGAATCCA 184  
7105 TTGACGCGAGAGATGAGAGGAAAGTATCCGTTCCGCGGAAATCTTCCGAGATCCA 7164  
185 AGAAATTTCCCAACGCGATGCCGATGCGGACGCCGAGTTAACAACCTCCGCTCGG 244  
7165 GGAATTTCCCGAGGATGCCATATGGGCAAGCCGAGTTACAACCTCCACTGATAG 7224  
245 AGTCCGGAAGGCGCCGAGTACTAGTCCCTCCATGTGATCTGGGTGCCACTGCCATTA 304  
7225 AGTCCGGAAGGCGCGGACTAGTACCTCCGATGATGAGGTGCCATTCGCACTTA 7284  
305 CTAAAGACCCCTCTATACCACTCCACGAGAAAGAGACAGTTGTTCTGACAGAAATCCA 364  
7285 CTAAAGACCCCTCTCAATACCACTCCACGAGAAAGAGAGGTTGTTCTGACAGAAATCCA 7344  
365 CCGTGTCTTCTGCGCTGCGGAGCTTGCCAACAAGGCTTTGGTAGCTCCGACCGTCCG 424  
7345 CCGTGTCTTCTGCGCTGCGGAGCTTGCCAACAAGGCTTTGGTAGCTCCGACCGTCCG 7404  
425 CCGTGCACAGCGGACGCGCAACCGCCCTCTGACCAATCTCCGACAGCGCGAGCAG 484  
7405 CCGTGCACAGCGGACGCGCAACCGCCCTCTGACCAAGCTCCGACAGCGCGAGCAG 7464  
485 GATCTGACGTTGATGATGATCTCTCATGCGCCCTTGAGGGGGAGCGCGGGAGCCCG 544  
7465 GGTCCGATGTTGATGATGATCTCTCATGCGCCCTTGAGGGGGAGCGCGGGAGATCCCG 7524  
545 ATCTCAGCGACGCGGCTTTGATCTTACCGTGTGAGAGAGCGCGGTGAGAGCTGTCTGCT 604  
7525 ATCTCAGCGACGCGGCTTTGATCTTACCGTGTGAGAGAGCGCGGTGAGAGCTGTCTGCT 7584  
605 GCTGCATGTCCTACATGAGACAGCGGCTCTGATCAAGCCATGCGCTGCGAGAGAAAGCA 664  
7585 GCTGCATGTCCTACATGAGACAGCGGCTCTGATCAAGCCATGCGCTGCGAGAGAAAGCA 7644  
665 AGCTGCCATCAACGCGGTTGAGCACTTTGCTGCGTCAACAATGATGTCTAAGCTTA 724  
7645 AGCTGCCATCAACGCGGTTGAGCACTTTGCTGCGTCAACAATGATGTCTAAGCTTA 7704  
725 CCAATATCCGCGACGCGACGCGCAGAGAAAGGTCACTTTGACAGACTGCAAAATCC 784  
7705 CAAATCTCGCAGCGCAGCAGCGCAGAGAAAGGTCACTTTGACAGACTGCAAGGTCC 7764  
785 TGGACGATCACTACAGAGAGCTCTCAAGAGATGAAAGCGGAGGCTCCAGTTAAG 844  
7765 TGGACGATCACTACAGAGAGCTCTCAAGAGATGAAAGCGGAGGCTCCAGTTAAG 7824  
845 CTAAAGCTTCTATAGTAAAGAGCTGCAAGGTGACGCGCCGACATTTGGGCAAAATCTTA 904  
7825 CTAAAGCTTCTATAGTAAAGAGCTGCAAGGTGACGCGCCGACATTTGGGCAAAATCTTA 7884  
905 AATTGGCTATGCGGAGAAAGAGCTGCGAACCCTATCCAGCAAGGCGATTTAACATATCC 964  
7885 AATTGGCTATGCGGAGAAAGAGCTGCGAACCCTATCCAGCAAGGCGATTTAACATATCC 7944  
965 GCTCCGTTGGAGAGACTTTGTTGAAAGACCTGAAACCAATTGACACACATCATATG 1024  
7945 TCTCCCGTTGGAGAGACTTTGTTGAAAGACCTGAAACCAATTGACACACATCATATG 8004  
1025 CAAAAAATGAGGTTTCTGCTGCTCAACAGAGAGAGGCGCAAGCCAGCTCGCTTA 1084  
8005 CAAAAAATGAGGTTTCTGCTGCTCAACAGAGAGAGGCGCAAGCCAGCTCGCTTA 8064  
1085 TCGTGTCCAGACTTGGGGGTCTGTGTGTCGAGAAATAGCCCTCTATGACGTGCT 1144  
8065 TCGTGTCCAGACTTGGGGGTCTGTGTGTCGAGAAATAGCCCTCTATGACGTGCT 8124

QY 1145 CCACCTCCCTCAGGCTGATGGGCTCTCGNAGAGATTCAGATATCTCTGACAGC 1204  
DB 8125 CCACCTTCTCTAGGCGCTGATGGGCTCTCTATACGGGTTCCAGTACTCTCTGACAGC 8184  
QY 1205 GGGTCAGATTCTGTGAAAGCGCTGGAATCAAGAGAGACCCCTATGAGGCTTTGATATG 1264  
DB 8185 GGGTCAGATTCTGTGAAATGCTGGAATCAAGAGAGAGACCCCTATGAGGCTTTGATATG 8244  
QY 1265 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATT 1324  
DB 8245 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATT 8304  
QY 1325 ATCAATGTTGATGATGAGCGCCCGGAGAGCGAGACGGCCATTAAGTGTCTCAAGAGCGGC 1384  
DB 8305 ACCAATGTTGATGATGAGCGCCCGGAGAGCGAGACGGCCATTAAGTGTCTCAAGAGCGGC 8364  
QY 1385 TTTATATCGGGGCTCCCTGACTAATTCAAAAGGCGAGATGCGGCTATCGCGGTGCC 1444  
DB 8365 TTTATATCGGGGCTCCCTGACTAATTCAAAAGGCGAGATGCGGCTATCGCGGTGCC 8424  
QY 1445 GCGCGAGCGGCTGCTGACGACTAGCTGCGTATACCTCAATGTTACTTGAGAGGCT 1504  
DB 8425 GCGCGAGCGGCTGCTGACGACTAGCTGCGTATACCTCAATGTTACTTGAGAGGCT 8484  
QY 1505 CTGACGCTGTGAGAGCTGAAAGCTCAAGACTGACAGATGCTCGTGTGCGAGAGCGGC 1564  
DB 8485 CTGACGCTGTGAGAGCTGAAAGCTCAAGACTGACAGATGCTCGTGTGCGAGAGCGGC 8544  
QY 1565 TTGTCTTATCTGTGAGAGCGCGGAAACCAAGAGAGCGCGGACTTACAGAGTCTTCA 1624  
DB 8545 TTGTCTTATCTGTGAGAGCGCGGAAACCAAGAGAGCGCGGACTTACAGAGTCTTCA 8604  
QY 1625 CCGAGGCTATGACTAGTACTCTGCCCCCGGGAGACCGGCCCAACAAGATATGAGCT 1684  
DB 8605 CCGAGGCTATGACTAGTACTCTGCCCCCGGGAGACCGGCCCAACAAGATATGAGCT 8664  
QY 1685 TGGAGTTGATACATATGATCTCTCAATGATGCGTGTGCGGACGATGATGCGCAAA 1744  
DB 8665 TGGAGTTGATACATATGATCTCTCAATGATGCGTGTGCGGACGATGATGCGCAAA 8724  
QY 1745 GGGTATACTACTCAACCCGCTGACCC 1769  
DB 8725 GGGTATACTACTCAACCCGCTGACCC 8749

RESULT 8  
AF208024 9373 bp RNA linear VRL 30-MAR-2000  
LOCUS Hepatitis C virus strain MD34, complete genome.  
DEFINITION AF208024  
ACCESSION AF208024  
VERSION AF208024.1 GI:7341102  
KEYWORDS  
SOURCE Hepatitis C virus  
ORGANISM Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.  
1 (bases 1 to 9373)  
Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and  
Sato, C.  
TITLE Characteristics of hepatitis C viral genome associated with disease  
progression  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 9373)  
AUTHORS Nagayama, K., Kurosaki, M., Enomoto, N., Izumi, N. and Sato, C.  
TITLE Direct Substitution  
JOURNAL Submitted (24-NOV-1999) Second Department of Internal Medicine,  
Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku,  
Tokyo 113-8519, Japan  
FEATURES  
source  
1..9373  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/strain="MD34"

gene  
CDS

```
/db_xref="taxon:11103"  
/note="from a patient with persistently normal ALT values  
[ASC-14]"  
1..9373  
/gene="MD34"  
330..9356  
/gene="MD34"  
/codon_start=1  
/product="polypeptide"  
/protein_id="AA061205.1"  
/db_xref="GI:7341103"  
/translation="MSINPKPQKTKRNTNRPDVKKPGSQIVGVYLIDRRPRL  
GVRATKTSERQPRGRQPIKARQPEBAAQPGYEMPLVGNMGWAGMLSPRG  
SRNPGTDPRRSRNLKVIDLTCGFADEMGYIPLVGAIPGAAARALAHVRLS  
GNYATGNLPGCSFSIPLIALISCLTIPASAEVRNVSCHYVITNDSSISLYEAD  
MIMSGCVPCYREGNVSRCWALPTLAAKRGSVPTTIRHVDLVGAAGFAMG  
VVDLGGSVFLVSQLPTSPRRHETVQDCNSYIPGHLSHRANMMNMNSTALV  
SOLLIPQAVMDVAGAHVLAGLAIYSWGNMAKVLVLLPFGVDMILIGSSQD  
TTRRTSLSPGSONIQLVKTEGMHINRTALNCDSHQTFALFPYAKINTSG  
LERLTSCHEIDKFDQGMGLTYEHGSDQKQCMWYAPRCGIPASVCGPVCF  
PSPVVGTTDRFGATYTWGENETVDLLNTRPQGMFGCTMNSSGYTKCGGP  
CNIGVGNDTLICPTDCPRKHPBATYTKGSGPMLTPRCMDVYPIRMHYPTVAF  
PKVRYVGVGVEHRLACNMTGRCNLPDRRSLSPLLSTTMOILPCTFLPA  
ISTGLIHQNIIVDQYLYGVSAVVSFTKMEYVLLPFLDLADATCAQMMPLIL  
QAEPLNLVNLMAASVAGAHGISLPLVFPCAWYIKKGLVGAAYALVGLPILLL  
LALPRAVAMDREMAASCGAVFVGLVLTLSPPYKFLARLIMSOYEITAEALQ  
VWVPLVNGRDAIILLTCAVHELIPDTIKLILGLPLVLAASIIKVPYPAQ  
GLIRACMLARKVAGHYVOMAFMLAALITGVVDHLPIDQMAASIRDLAIVAE  
VSDMETKIIWGAADTAACGDIILGLPYARPKKELILGDSQCMRLATAY  
SQTRGLIGCTITSLTGRDKRNGEVOVSTATOSPLATCNGCTVCFHAGSKTL  
AGKQPIYQMTYVNDQDLVWQAPPGASLITCTGSSDLYLVTHADIVPRRSGTS  
RGLSLPRVYSLVKSGSGPLCPSGHAGVIFRAVCTRGVAKAVFVPEMETMR  
SEVFDNSPPVAPQTPVAAHLAPTSQSKSTKPAAYAAQYKLVINPVAALTF  
GAYMKAGVDPNIRTGARTTGAIPITYSTYKFLADGCGGAYVIMDECHSD  
STSIIGTVDLDAETAGARLVLAATPQSVTVPHPIEBVALSNGEIPPKKAI  
PIETIKGRHLIFCHSKKCDLATKLSGLINNAVYRGDEVSTPAGNVVATD  
ALMTGIGDPSVIDCNTCVITVPSLDPITLITTTVPQDAVSQRKGTGRR  
GIYREVTPGEPSGMDSSVLCEDYAGCAWELPAETVSRLAYLTPGLPVCODI  
LEFWSVETGLTHIDAHFLSQKQADNPYLAQATCAQAPSPMDMCLIL  
RLKPTLHGPTPLRLGAVONETLTHPTKTYIMACMSDELVTSTYQGLVLAAL  
AAVCLTGSVVIVGRILISGRPAVLPDEBVLREFPEDECHSLPYIEQMGLEOF  
KOKAGLIOTAKQAVAPVYESKRALBAFMAKMMNFISGITYLAGSLTPENPA  
IASLAFDASITSLTOSTLFLNLDGVAALAPPSALSFVAGAGAAVSGIL  
GKVLVIDLAGYAGVALVAFKNGSEMPSTEDLVNLPALISGALVAVGVA  
RHVPGEGAVQMMARLIAFASRGHNSPTHVPESDAPLPTOLSSITTCOLKRL  
HOWINDECTPCSGMLRBDMDICVLDFTMLOSLKPLPLPVPSPSCORGYK  
WEGDGIOTTCGCAOITGCHVNSMRILIGPTCSNTWAGTPINAYTTGPTSPAP  
NYSRALMRYAAEEYVTRVGDPHYVTGTDNVKCPQVPAEPFTFVDRKRLYA  
PCKPLREEVTFVGLNOYLVSQLPCPEPDVAVLTSMLTDPSTHITAEAKRLAR  
GSPSIASSASQLSAPSLKATCTTHSDPADLIEANLMPDEKGNITRVESNKY  
VLDSFDPLRAEDREVSVAELIRKSKFPAMPVAPYNNPILSMDPDYVP  
PVHGCPLPTKAPPIPPRRKRTVLVETSSVALAELATPTSPSSSADSDGAT  
APDPSDMSDAGSVESYSSMPLLEGGBDPLSDGMSYTSBASADVYCCNSYT  
WTGALITPCAAEGSKLPIVALSNSILRHNNYATSSASQRQKSVFDRLOVLDH  
YRVILKEMAKASIVAKILVSEACKLPKPSASKRGYAKDVRNLSGKAVNIRS  
VKDLLEDETETIDITTIKAKNEIFCVQEPKGRKPARLIVFEDLVRNICEKALNIRS  
STLPQAWASXYGQYSGQVLEFLVAMKSKCMGPAVRCDSVITYEDJIVER  
SYCCDCLAPAEARQARSILTERLYVGLPINSKQNGCRRCRAGVLTSGNLTLC  
YIKASACRAKLRDCTMLVCGDDLVICESAGQEDASLAPFEATMRTPYASPGDP  
PQPERYDELITSCSSNVSVAHDASGRYVYLTDRPTPLARAMEFTASTPYNSP  
LITVAPTLMARMLMTWTFPSILAOQEKALDCQIYACVSIIEPLDIPOIQRHLG  
SAFSLHYSYSGEINRVASCLRLGVPLVWRHARSYAKLLSGGGAALCGKLFN  
MAVRTKLKLTPIPAASOLDLSNMVFAVSGSDIYHLSRARRRMFLCILLISVGI  
YLPNR"
```

## ORIGIN

Query Match 88.7%; Score 1570.6; DB 14; Length 9373;  
Best Local Similarity 93.0%; Pred. No. 0;  
Matches 1645; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1 CAAATGACTTCCCGAGAGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATAGATG 60  
DB 6975 CATCATGACTCCCGAGAGCTGACCTCATCGAGGCCAACCTCTGTGGCGGAGAGATG 7034

```
QY 61 GACGGGACATTACCCGGTGGAGTCAGAGAACAAAGTAGTAATCTGGAATCTTTCGAC 120  
DB 7035 GCGGAAACATCACCCGGGTGGAGTCAGAGAAATAGTAATCTGGAATCTTTCGAC 7094  
QY 121 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTTCGGAAA 180  
DB 7095 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCGGCGGAGATCTTCGGAAA 7154  
QY 181 TCCAGAAATTCACACGAGCATGCCGATGGGACGCCCGGANTTAACCTCCGCTG 240  
DB 7155 TCACAGAAATTCACCCCGGAGATGCAVATGGGACGCCCGGANTTAACCTCCGCTG 7214  
QY 241 CTGAGCTCTGAAAGAGCCCGGAGACTACGTCCTCCAGTGTATACGAGTGCCTGCGCA 300  
DB 7215 CTAGAGCTCTGAAAGAGCCCGGAGACTACGTCCTCCAGTGTATACGAGTGCCTGCGCA 7274  
QY 301 CCTACTAAGACCCCTCTTATACACCTTCACGAGAAAGAGACAGTTGTTCTGACAGAA 360  
DB 7275 CCTACTAAGACCCCTCTTATACACCTTCACGAGAAAGAGACAGTTGTTCTGACAGAG 7334  
QY 361 TCCACCGTCTTCTGCGCGGAGGCTTCCACAAAGCTTTGTAGTCTCGGAGCG 420  
DB 7335 TCCACCGTCTTCTGCGCGGAGGCTTCCACAAAGCTTTGTAGTCTCGGAGTGC 7394  
QY 421 TCGGCGTCCGACAGGCGACCGGCAACCGCCCTCTGACCAATCTCCGACGAGCGCGA 480  
DB 7395 TCGGCGTCCGACAGGCGACCGGCAACCGCCCTCTGACCAAGCTTCGACAGTGC 7454  
QY 481 GCGAGATCGACGTTGAGTGTATATCTCTCATGTCGCCCTTTGAGGGGAGCCGGGGAC 540  
DB 7455 GCAGAGATCGACGTTGAGTGTATATCTCTCATGTCGCCCTTTGAGGGGAGCCGGGGAT 7514  
QY 541 CCCGATCTCAGCGAGCGGCTTTGGCTTACCGTGAAGTGAAGGCGCGAGAGAGCGTGC 600  
DB 7515 CCCGATCTCAGCGAGCGGCTTTGGCTTACCGTGAAGTGAAGGCGCGAGAGAGCGTGC 7574  
QY 601 TGTGCTGATGTCCTTACATGAGCAAGGCGCTGTATCAACGCAATGCGTCCGAGGAA 660  
DB 7575 TGTGCTGATGTCCTTACATGAGCAAGGCGCGCTGTATCAACGCAATGCGTCCGAGGAA 7634  
QY 661 AGCAAGCTCCCATCAACGCGTTGAGCAATCTTTGCTGCTCACACACATGCTTAC 720  
DB 7635 AGCAAGCTCCCATCAACGCGTTGAGCAATCTTTGCTGCTCACACATGCTTAT 7694  
QY 721 GCTACCATATCCGAGCGGAGCGACGCGAGGAGAAAGTCACTTTGACAGACTGAA 780  
DB 7695 GCTACCATATCCGAGCGGAGCGACGCGAGGAGAAAGTCACTTTGACAGACTGAA 7754  
QY 781 ATCTGAGCGATCACTACAGAGAGTGTCTCAAGAGATGAAGGCGAGCGTCCACAGTT 840  
DB 7755 ATCTGAGCGAGCATCACTACAGAGAGTGTCTCAAGAGATGAAGGCGAGCGTCCACAGTT 7814  
QY 841 AAGGCTAAGCTTCTATCATGTAAGAGAAAGCTTGCAAGCTGACGCCCCCACTTTCGCAAA 900  
DB 7815 AAGGCTAAGCTTCTATCATGTAAGAGAAAGCTTGCAAGCTGACGCCCCCACTTTCGCAAA 7874  
QY 901 TCTAAATTTGGCTATGGGGCAAAAGAGTCCGGAACCTATCCAGAGGCGCATTAACAC 960  
DB 7875 TCTAAATTTGGCTATGGGGCAAAAGAGTCCGGAACCTATCCAGAGGCGCATTAACAC 7934  
QY 961 ATCCGCTCCGTTGGGAGAGCTTGTGGAGACACTGAAACACCAATTGACACCACTC 1020  
DB 7935 ATCCGCTCCGTTGGGAGAGCTTGTGGAGACACTGAAACACCAATTGACACCACTC 7994  
QY 1021 ATGCAAAAAATGAGGTTTCTGCGTCCACCAAGAGAGAGCGCCGACAGCTGCG 1080  
DB 7995 ATGCAAAAAATGAGGTTTCTGCGTCCACCAAGAGAGAGCGCCGACAGCTGCG 8054  
QY 1081 CTTATCGTTCCTCAGACTTGGGGTCCGCTGTCGAGGAAATGCGCCCTTATGACGTG 1140  
DB 8055 CTTATCGTTCCTCAGACTTGGGGTCCGCTGTCGAGGAAATGCGCCCTTATGAGTGTG 8114
```



OY		1141	GCTCCACCCTCCCTCAGCGTGTGATGGGGCTTCCTGTACGGATTCCAGATTATTCCTCGGA	1200
Db		8115	GTTCTCACACCTTCTCTCAGGCCGTGAWG93CTCTCATACGAAATTCAGTAATCTTCCTCGGA	8174
OY		1201	CAGCGGAGTGAGTTCCTCGGTGMAACCCCTGSAATAATCAAGAAGAACCCCTATGGGCTTTGCA	1260
Db		8175	CAGCGGGTTGAGTTCCTCGGTGMAATCCCTGSAATAAAGAAATGTCTTATGGGCTTCGCA	8234
OY		1261	TATGACACCCCGCTGTTTTGACTCAAAGTCACTGAAATGACATCCGTGAGAGAGTCA	1320
Db		8235	TATGACACCCCGCTGTTTTGACTCAAACGCTCACCGAGATGACATCCGTGAGAGAGTCA	8294
OY		1321	ATTATCAATATGTGACTTGGCCCCCGGAACGCACAGGCCATPAAGTCCCTCAACAG	1380
Db		8295	ATTATCAATATGTGACTTGGCCCCCGGAAGCCACAAGGCCATPAAGTCCCTCAACAG	8354
OY		1381	CGGCTTTATATCGGGGGTCCCCTGACTAATTCAAAAAGGAGAACTGCGGCTATCGCCGG	1440
Db		8355	CGGCTTTATATCGGGGGTCCCCTGACTAATTCAAAAAGGAGAACTGCGGCTATCGCCGG	8414
OY		1441	TGCGCGCGAGCGGCGTGCCTGACGACTAGCTGCGGTAAATACCTTCATGTTACTTGAAG	1500
Db		8415	TGCGCGCGAGCGGCGTGCCTGACGACGAGCTGCGGCAACACCTTCATGTTACTTGAAG	8474
OY		1501	GCTTGTGACGCTGTCTGAGCTGCACAAAGCTCCAGAGCTCAAGATGCTGTGTGGGAGAC	1560
Db		8475	GCTTGTGGGCTGTCTGAGCTGCAGAGCTCCGGGACTGCAAGATGCTGTGTGGGAGAC	8534
OY		1561	GGCCTTGTCGTATATGTGTAGAGCGCGGGAAACCCAGAGAGACGGGGGAGGCTPACGAGTC	1620
Db		8535	GACCTTCGTCTTATTTGTGTAAAGCGGGGAAACCCAGAGAGATGGGGCAGGCTACGAGTC	8594
OY		1621	TTTACGAGAGCTATGACTTAGTACTGTGCCCCCCCCCGGGAGACCGGCCCAACAGAAATAC	1680
Db		8595	TTTACGAGAGGCTATGACTTAGTAAAGCGGGGAAACCCAGAGAGATGGGGCAGGCTACGAGTC	8654
OY		1681	GACCTGGAAGTTGATTAACATCATGCTCTCCAATGTGTGTCGGCGACAGATGACTTGGC	1740
Db		8655	GACCTGGAAGTTGATTAACATCATGCTCTCCCAACGTGTGATCGGACACAGATGACTTGGC	8714
OY		1741	AAAAAGGATATCTACTACCTCACCCCGTAGACC	1769
Db		8715	AAAAAGGTACTATCTCAACCCGTGACC	8743
RESULT 9				
AY587845		9357 bp	RNA	linear
LOCUS	Hepatitis C virus strain RFL_2k/1b, N687 polyprotein gene, complete cde.	VRL 25-JUN-2004		
DEFINITION	AY587845			
ACCESSION	AY587845			
VERSION	AY587845.1	GI:46560635		
KEYWORDS				
SOURCE	Hepatitis C virus			
ORGANISM	Hepatitis C virus			
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			
REFERENCE	1 (bases 1 to 9357)			
AUTHORS	Kalilina,O., Norder,H. and Magnus,L.O.			
TITLE	Full-length open reading frame of a recombinant hepatitis C virus strain from St Petersburg: proposed mechanism for its formation			
JOURNAL	J. Gen. Virol. 85 (Pt 7), 1851-1857 (2004)			
PUBMED	15218169			
REFERENCE	2 (bases 1 to 9357)			
AUTHORS	Kalilina,O., Norder,H. and Magnus,L.O.			
TITLE	Direct Submission			
JOURNAL	Submitted (01-APR-2004) Virology, Swedish Institute for Infectious Disease Control, 171 82 Solna, Sweden			
FEATURES	Location/Qualifiers			
source	1..9357			
	/organism="Hepatitis C virus"			
	/mol_type="genomic RNA"			
	/strain="RFL_2k/1b_N687"			

```

5'UTR
CDS
/db_xref="taxon:11103".
<!. "natural recombinant"
1..268
269..9313
/note="encodes C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, NS5b"
/codon_start=1
/product="polyprotein"
/protein_id="AAR00644.1"
/db_xref="GI:4656036"
/translation="MSTNPKPQPKTKRNTNRPODVKPRGGGQIVGVVLIILPRGRGVRNTPRSGSRSGPGRGRORIPIKDRSRSGWGPGYWMYKGNBICGMALISPRSRSPWGPTDPRHRSNRNGKVIDITTCGPAIDMGITPVGAVYGYARLALAGRVLEGINATNGIPGCSFIFLLALSLCSIPVSAVEVKNKTSQTMATNDSNSITWLEAVLHVPGCIPEELVGNVSRMWIPVTPNVAVEREALTKGLTATIDLVASTALIGDVGCAAMIQAQAVIISPERHAFVQDNCISIVGHITGHMMADNMWPTTMMVAYLKRIPETVIDITGAMWFGVLAFYSCMGAFVLIILLDAGVASTITWGGQSHTMAGLVSIFSGPARONIDLIINTNGNINIRTLNCDISLQSLPVPHPNSGCCPRMSCRPIEGEPRFRIQMSLKYENVTNDEMRCYCMHPEYPCGIVPAPYCPGYCPSPVAVVGTTDKYGAPYTWBENEDVFLNSTPRPAMFGCTMMNQTGTYKTGPAPCIRIPEDNSTEDILCPDCEPKRKEAAYTTCGAGRPVLPCLVDYPRALHYPVGNPTPIKRIEMFICGVBRHLNAAENCFRGGDNCENRDSLSLSTHTWALIPCSLTHWALYSDNPALSTGLIHLHONIVDQYLYGSPATKTRIKVMEVYLIPLILADRYACMLMLLIILGOBALKEVLILHAASAPNGISICFTTPIPAHHIKRVVPLATVYGLKMSFSLILALPQOAYALNITBEGQLGVLLVIVSLTSPAIKILCSLMLSTLVLAALAIQEWBPLOARGRGDGIIMAAVILCPQIVDITKMLVALGPHLARSVLTVALEVAOALRLICAVRHLSGGKYQWMLTLGMRITGYIYHLPIDMAHAGRLDILVAVPVEESDMETKITMGATETACGQIILGLPSARGREIILGADSLSGQGRLLAPIPAVSQDREGILGIVTLTGDRNOIVEGVNVSATQSPLATQNGVCMVYHAGSKTLAPKPGPIITQNTNDODLVGNPAPGASLTPTCGSSDLVITRHADVIPRRDRDSRGLSTSPRISVLYKSGSGPLICPSGNAVIGIPRAVCTRGVAKADVAVESMETTRSPVPTDNSPPAVPQTFQVAAHLAPTGSGSKGPPAAVAAQYKVLVNPSSAATGFGAYMSKAGVDPNRTGRTITTGAPITVSTYGTKLADGGSGGADIDIMCTECHSTDSILIGITVLDQAEATAGRLVLTATPPGSAVPHPSIEEVLNSNGEIIYVGAIPLEVYKGGRLIFCSGKCDKELAKISLGLNAYYRGLDVSYPIPSGDVVAVALDMGYTDEPDSVIDCNTCTQTVDFDSLDPDTFTLETTMPDQAVSRQGRGTGRRGGIRFVYVGERSPGMFDSVCECYDACQAVELTPAETSVRLAYLNPGLPVCDHLLEPESVPTGLTHIDAPLSQTKQGENPFLYIAQATVCARQAPSPMDVMKCLIRLKEPTLHPTPLVRLGAVNCEVLTLPITFELIMACMDLEVTSTVLYVGVLAVALCJTSSVYIGRIVLSGPAIIPDEBALYOEDMEBCASHPIYEOGQLOAEQFOKALGILHTPAKQAEAAAPVESKWPALETPMAGMNPISGLOYAGLSLIPGNPAIASLMAFTASTSPITTOHTLIFNLGCVAAQLAPSAASAFDAGIAGCAVSGIGLKEVLDIAGYGAGVAGALVAFKMSPTBEDVNLNPLISGLALVGVGCAILRRHVGKGVAVQVMNRLIAPASRGHVSPTBEDVNLNPLISGLALVGVGQLERLHONINEDSTPCSGSMLDWDMLCTCTVLTDEKTMOSLILPOLPEVPFSCPGRYGQWVRGDQIMQTTCPGKQOITGHYKNSMRIVGKTSGMTHGPEIYAAITPGPSPALNYSRALRVAAEVYEVTRVDPHYVGMITDNKCPQVAPREFTEVDGRLHRYADACKPRLREEVTFQVGLQNYIVGSQLCEPREDVAVLISMLOPESHITBETRIARLARGSPSLASSSASQSLAPSLKATCTTHDSPADILLEANLWROEGNITRSEBKNPVLDSFPIREEDBEREVSVAELILRSKRSPAMPIARPDYNPFLSSMAADPDVHPVHGCLPPTKAPRIPEPRKRPVTLRESVSAALAIATLPTFGSSSAAVCSGTVYKLSAASACRAKARDCTMLVCGDLVYICSAQOEDASIRARVEAMTRYSAPREDPPQPEYDLELITSCSNNVSVAHDAGKRYIYTRDPTTILARAAMETANTPNSWIGNTIMTAFTPLMAMLIIMTHFSLILAOEOLKADCOIYACVSIIEPLDLPQIILORHLGAFSLHSYSPEGINRVASCLKRLPPIRARHRRVSRVAKLISGGAGATGKLYLFNMAVYTKLKLTPIPASQLDLSGMEFVAGSGDIYHSISRARBRPMWCLLILSVGIVYLIPNR"
9314..>9357

```

QY 61 GGCGGGAGCATTAACCGCGTGAAGTCAAGAA CAAGTAGTAATCTTGGACTCTTTGCAC 120  
 Db 6992 GGGGGAATACATACCGCGGTGAGTCAAGAAATAGTAGTAGTTTGGACTCTTTTGAC 7051  
 QY 121 CGGCTCGAGCGGAGGAGATGAGCGGAAGTGTCCGTCCGCGGAGATCTTGGGAAA 180  
 Db 7052 CGGATCCGCGCGGAGAGGAGCAGAGAGGAAGTATCCGTTCGCGGAGATCTTGGCGAAA 7111  
 QY 181 TCCAAAGAAATTCACACAGCGAATGCCGAGATGGGGACCGCGGATTACACCTCCGCG 240  
 Db 7112 TCCAGAAATTCGCCAGAGATGCCATATGGGACCGGCGGATTACACCTCCGCG 7171  
 QY 241 CTGAGATCTTGAAGGCGCGGAGCTACGTCCTCCAGATGTAATGGTGGCCACTGCA 300  
 Db 7172 TTAGGCTCTTGAAGGAGCCCGGACTACGTCATCCGTGTGTGACGCGGTCCTGCA 7231  
 QY 301 CCTACTAAGACCCCTCTTATACCACTTCAACGAGAGAGAGACAGTTGTCTGACAGAA 360  
 Db 7232 CCTACCAAGGCGCCCTCAATACCACTTCAACGAGAGAGAGACAGTTGTCTGACAGAG 7291  
 QY 361 TCCACCGTGTCTTCCGCGCGGAGCTTGCACAAAGCTTTGTGTAGTCCGAGCCG 420  
 Db 7292 TCCACCGTGTCTTCCGCGCGGAGCTTGCACAAAGCTTTGTGTAGTCCGAGATG 7351  
 QY 421 TCGGCGGTGAGAGCGGAGCGGAGCGGCGCTCTGACCAATCTCCGAGAGCGGAG 480  
 Db 7352 TCGGCGGTGAGAGCGGAGCGGAGCGGCGCTCTGACCAATCTCCGAGAGCGGAG 7411  
 QY 481 GCAAGATCTGAGCGTGTGATGTTCTTCATGCCCCCTTTGAGGGAGAGCGGAGAG 540  
 Db 7412 AAGAGATCTGAGCGTGTGATGTTCTTCATGCCCCCTTTGAGGGAGAGCGGAGAG 7471  
 QY 541 CCGATCTGAGAGAGCGGAGCTTGTGTACCGTGAAGAGAGCGGAGAGAGAGAGAG 600  
 Db 7472 CCGATCTGAGAGAGCGGAGCTTGTGTACCGTGAAGAGAGAGAGAGAGAGAGAG 7531  
 QY 601 TCGTGTGATGTCTTACATGAGAGAGCGGAGCTTGTATGACGCAATGCGTGGAGAA 660  
 Db 7532 TCGTGTGATGTCTTACATGAGAGAGCGGAGCTTGTATGACGCGTGGAGAGAA 7591  
 QY 661 AGCAAGCTGCCATCAAGCGGTTGAGCAACTCTTGTCTGCGTCAACAAATGCTTAC 720  
 Db 7592 AGCAAGCTGCCATCAAGCGGTTGAGCAACTCTTGTCTGCGTCAACAAATGCTTAT 7651  
 QY 721 GCTACCAATCCGAGCGGAGCGGAGCGGAGAGAGAGTCACTTTGACAGACTGCA 780  
 Db 7652 GCTACCAATCCGAGCGGAGCGGAGCGGAGAGAGAGTCACTTTGACAGACTGCA 7711  
 QY 781 ATCTTGAAGATCACTACAGAGAGCTGTCAAGAGATGAGCGAGAGCGTCAACAGT 840  
 Db 7712 GTCTTGAAGATCACTACAGAGAGCTGTCAAGAGATGAGAGAGCGTCAACAGT 7771  
 QY 841 AAGGCTAAGCTTCTATCAGTAGAGAGAGCTTGAAGCTGACGCGCCCACTTGGCCAA 900  
 Db 7772 AAGGCTAAGCTTCTATCAGTAGAGAGAGCTTGAAGCTGACGCGCCCACTTGGCCAA 7831  
 QY 901 TCTAAATTTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 Db 7832 TCTAAATTTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7891  
 QY 961 ATCCGCTCCGTGTGAG 1020  
 Db 7892 ATCCGCTCCGTGTGAG 7951  
 QY 1021 AAGGCAAAATGAGAGTTTCTGCGTCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
 Db 7952 AAGGCAAAATGAGAGTTTCTGCGTCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 8011  
 QY 1081 CTATATGTTTCCAG 1140  
 Db 8012 CTATATGTTTCCAG 8071

QY 1141 GTCTCCACCTCCCTCAGAGCTGTGATGGGCTCTCTGTACGAGATTCAGATATCTCTGGA 1200  
 Db 8072 GTCTCCACCTCCCTCAGAGCTGTGATGGGCTCTCTGTACGAGATTCAGATATCTCTGGA 8131  
 QY 1201 CAGCGGCTGAGGTTCTGTGTGAAGCGCTGTGAATCAAGAGAGAGAGAGAGAGAGAG 1260  
 Db 8132 CAGCGGCTGAGGTTCTGTGTGAAGCGCTGTGAATCAAGAGAGAGAGAGAGAGAGAG 8191  
 QY 1261 TATGACACCGCTGTTTGAATCAACAGTCACTGAGATGATCATCCGTGTGAGAGATCA 1320  
 Db 8192 TATGACACCGCTGTTTGAATCAACAGTCACTGAGATGATCATCCGTGTGAGAGATCA 8251  
 QY 1321 ATTATCAATGTGTGATCTTGGGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
 Db 8252 ATTATCAATGTGTGATCTTGGGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8311  
 QY 1381 CGGCTTATATCGGGGCTCCCTGACTATATCAAGAGAGAGAGAGAGAGAGAGAGAG 1440  
 Db 8312 CGGCTTATATCGGGGCTCCCTGACTATATCAAGAGAGAGAGAGAGAGAGAGAGAG 8371  
 QY 1441 TGGCGGAGAGCGGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
 Db 8372 TGGCGGAGAGCGGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8431  
 QY 1501 GCTCTGAGAGCTGTGAG 1560  
 Db 8432 GCTCTGAGAGCTGTGAG 8491  
 QY 1561 GGCCTTGTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
 Db 8492 GGCCTTGTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8551  
 QY 1621 TTACAGAGAGCTATGAG 1680  
 Db 8552 TTACAGAGAGCTATGAG 8611  
 QY 1681 GACCTGAGGTTATTAATCAATGCTCTCAATGTGTGTGAGAGAGAGAGAGAGAGAG 1740  
 Db 8612 GACCTGAGGTTATTAATCAATGCTCTCAATGTGTGTGAGAGAGAGAGAGAGAGAG 8671  
 QY 1741 AAAAGGTTATTAATCAATGCTCTCAATGCTGTGAGAGAGAGAGAGAGAGAGAG 1769  
 Db 8672 AAAAGGTTATTAATCAATGCTCTCAATGCTGTGAGAGAGAGAGAGAGAGAGAG 8700

RESULT 10  
 HECKIR2  
 LOCUS Hepatitis C virus (strain HCV-1b, clone HCV-K1-R2), complete genome sequence.  
 DEFINITION  
 ACCESSION D50481.1 GI.1030705  
 VERSION  
 KEYWORDS Polyprotein; core protein; E1; E2; NS2; NS3; NS4A; NS4B; NS5A; NS5B; envelope protein; non-structural protein; interferon-sensitive; interferon-resistant; IFN-sensitive; IFN-resistant; ISDR; interferon sensitivity determining region; HVR; hypervariable region.  
 SOURCE Hepatitis C virus  
 ORGANISM Hepatitis C virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
 REFERENCE 1 (sites)  
 AUTHORS Enomoto,N., Sakuma,I., Asahina,Y., Kurosaki,M., Murakami,T., Yamamoto,C., Izumi,N., Maruno,F. and Sato,C.  
 TITLE Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region  
 JOURNAL J. Clin. Invest. 96 (1), 224-230 (1995)  
 MEDLINE 95340824  
 PUBMED 7542279  
 REFERENCE 2 (bases 1 to 9410)  
 AUTHORS Enomoto,N.  
 JOURNAL Unpublished

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

3 (bases 1 to 9410)  
Enomoto, N.  
Direct Submission  
Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental  
University, Second Department of Internal Medicine, 1-5-45 Yushima,  
Bunkyo-ku, Tokyo 113, Japan (E-mail: PYN04522@niftyserve.or.jp,  
Tel: 03-3813-6111 (ex.3224), Fax: 03-3818-7177)

FEATURES  
source

CDS

1.9410  
Location/Qualifiers  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/strain="HCV-1b"  
/db\_xref="taxon:11103"  
/clone="HCV-K1-R2"  
/note="IFN-resistant HCV from patient 2"  
330..9362  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="BAA09072.1"  
/db\_xref="GI:1814086"  
/translation="MSTNPKQKTKRANTNRNPDVYKFGGGQIVGGYVLLPRGRPL  
GVRAIRKTSERSQPRKQPIPAQRPQEGMAAQPEYPMPLVGENEGMAGWLLSRG  
SRSPMGPTDPRSRSMAGKVIDTLTCGADLMGVIPLVGAPLGASRALAGRVLESD  
GNVATGNLPGCSFSIFLLALSLCTIPASAEVRASGVHTVNDSCNASIYEAD  
MIMHTGCVPCVREANSRCWVALPTLPAARNSVPTTIRHVLLVGAARFCSAMY  
VGDLCGSVFLVSQLPTSPRRHETVDDCSISYRGVSGHRAMAMMMMMSPPTALV  
SOLIRIPQAVMDVAGAHGVTLAGLAYSIVGMAKAVLLVMLPAVDGNTYVSGAK  
SHTQGLVSLFAPGAQOKIQLVNTNGSMHINRTALCNDLSLNTGLALPYTHRFSS  
GCEPRLASCRSIDAFAGWGPIYVABGSDORPCMHYAPRCQIVPASEVCGPYC  
FTPSPVVGTDRSGAPITVSGENETVSLNNTAPQGWECCTMMNGTGFCTGCG  
PCNIGVGNNLTCTPCDFRKHBAATYKCGSPMLTPTRCMDYDYLHNYCTVNF  
TIFKVMVYGVGEHRMLAAMNTRGERCELDREBSLPLILSTENMOLIPSPFTLV  
PIVSTGLHILHONIVDVTLYGSAVSVFYIKETILLFILADARCAQCIAMMLL  
IAQAEALENLVNLAASVAGAGFLSPVFCAAVYIKGLVPAVAAYFYGWELL  
LIALPRAVYAMDREMAASCGAVFGALLTSPHYKVLVALIMLQYFTIRACAP  
AQLWIPLNVRGGRDAILLTCAHPELIFSTIKILVALIGLPMVAQGITPKPYFR  
ACGLIAPACMLVRVAGAHVYOMAFMLALITGVYVDHLPTLDMATGRLDAVVE  
PVESDMETKIITWGDTAACGDIILGLVASARKEITPCTGSESGSGGAAIAT  
AYSQTRGLGCTITSLTGRDKQVEGEVQVSTANOSFELATCTTGCCTVYHGAESK  
TLAAGKPIITQWYTNVDQLVGMQAPSGARSLPCTCGSSDLVLRHADVILPVRRG  
DGRGSLSPRPVSYLVSSGGPLCPGSAHVGIFRAVCTVAGAKAVDPIPEMSMTT  
MSPVFTDNSSPPAVPQTFQVAHLHAPTSKGNKPYEVAAGYKVLVNLSPVATL  
GFGVMSKAGVDNIRTCGVRTITGGAPIITYSYGKFLADGGSGGAGYIILICDEBGS  
TGSTSLIGITVDQAEETAGARLVYATPTPESYVVRHPIEVALNSTGERIPGK  
ALPITIKGRHLIFCHSKKCCDELAUKLSGGINAVYRGIDVSVIPASGVVVVA  
TQALMTFTGDEPDSVIDCNTCVTQVDFSLDPTLETITVTPDPAVRSQRRRTGRG  
RAGIYRFVTPGERSPSGMPSVLCCEYDGCAMVELTPAETSRILRAYINTPGLPVWQ  
DLHEFVESYFTGLTIDAHPLSOTKOAGNFPYLVAQATVCABAQAPPSPMDQMKC  
LIRKPTLHGPPLIYRLGAVONEVILTHPTIKYIACMSADILEVTSMTLVGVLAA  
ALAAVCLTGSVVIYGRITILSKRPALIPREVLVYREFDEMEBASHLPTIBQMOLAE  
QPRQKAGLLQRTAKQEAAPVRESKQALLETFAKHMNPLISGQYLAGLSTLFGN  
PAIASLMAFTASTISPLTQHTLLFNILGEMVAAQLAPPAASAFQARVAGIAGAVSI  
GAGKVLVDILAGAGVAGALVAFKVMGSGMPSTEDLVLLPLILSPGAVGVCAA  
ILRBHVPEGAVOMNRLIAFASRGHNSPTHVPESSAAAVTOLILSLITVTLK  
RLHOWINEDCSTPSCGSMLRDWDMTCTYLSDEKTMLOGRVLPRLPGVFLSQGRYK  
GIMRGDGINOTTCPGCAQITGHVYKMSMIVGPKTSNTHGTFPINAITTGCTSP  
APNYSALMRVAAEYEVTRVGDVHYVGMITDNLCPCQVPAFFEFELDEGRILHR  
YAPASKPRLRDEVTQVNLQYVVGSLCCEPREDVAIVTSMLTDSHLYTAETKARL  
AAGSPSLASSSSQLASPLKATCTTHDSPPADILEANLROMCGNITVESEN  
KVIILDSFEPRLAEDEREVSILPAELIRSKRPRAMPILARDVNPPLIESKNDY  
VPVYVGCPLPTKAPRIPPRKRTIVLTSTSVSALAEIKATKTPSSSSAASDGT  
ATAAPPOSSDDAGSDVESYMPLEBEGDPDLSDSKMSVSEBAEDVYCCMS  
YMTGLALTPCAEBSKPIINALSNSLRLHNNHNAVITRSVARSQRRKVTFDLOVD  
DHRYDLKEMKAKASTVKARLLSVEACKLTLPHSARSKFVGARQKVRNLSLAINH  
RVMXKDLQEDTEPTIDTITMAKNEVFCVQPEKGRAPALVYEPDGVCEEMALYD  
VSTLPOAVGASVGYOYSGORVEFLVNAWSKKCPMGAVYTRCFDSVLENDLRV  
ESISYOCCLAPBARQIRSLTERLYIGLPTIRYKCCONCGYRACSGVLTLSGNTL  
TCYKATACRAKAKODCTMLVCGDVLVYICBSAGTQEDBASIRVTRBAMTRSPAPG  
DPEQPTDELELITSCSNVSAHADSGKVVYILTRDPTPLAANAEHTAHTVNSAPL  
GNIMVAPYILMAAMILMTHFSLILAQEOLEKALDOQIGACCAVLPQIIOQLH  
GISAFSLHYSRBEINRVASCLRKLGVPVLRVHRARSRACLLSQGRAPALICGYTL  
FMAVARTKLTLPPIPAASQDLISGFMVAGSGDIYHVSRAPRFWMMLLLLSVGV  
GYLLPNR"

	mat_peptide	330..902	/product="core protein"
	mat_peptide	903..1478	/product="E1"
	mat_peptide	1479..2756	/product="E2"
	mat_peptide	2757..3407	/product="NS2"
	mat_peptide	3408..5300	/product="NS3"
	mat_peptide	5301..5462	/product="NS4A"
	mat_peptide	5463..6245	/product="NS4B"
	mat_peptide	6246..7586	/product="NS5A"
	mat_peptide	7587..9359	/product="NS5B"
	mat_peptide	1479..1559	/note="hypervariable region (HVR)"
	misc_feature	6354..7073	/note="interferon sensitivity determining region (ISDR)"
	misc_feature		
ORIGIN			
Query Match	88.3%;	Score 1562.6;	DB 14; Length 9410;
Best Local Similarity	92.7%;	Pred. No. 0;	
Matches 1640;	Conservative	0;	Mismatches 129; Indels 0; Gaps 0;
QY	1	CAAAATGACTTCCAGACGCTGACTCATGAGGCCAACCTCTGTGGCGGATGAGATG	60
Db	6981	CATCATGACTCCCGACAGCCCGACCTCATGAGGCCAACCTCTGTGGCGGAGATG	7040
QY	61	GGCGGGGAATTAACCGCGCTGAGTCAAGAACCAAGTAACTTCTGACTTTTCGAC	120
Db	7041	GGCGGGGAATTAACCGCGCTGAGTCAAGAACCAAGTAACTTCTGACTTTTCGAG	7100
QY	121	CCGCTCGAGCGGAGGAGATGAGCGGGAAGTCCGTCGCGCGGAGATCTCGGAA	180
Db	7101	CCGCTTCGAGCGGAGGAGATGAGCGGAGTATCCTTCGCGGAGATCTCGGAA	7160
QY	181	TCCAGAAATTTCCACACAGCATGCCCCGATGGGCAAGCCGGAATTACAACCTCGCTG	240
Db	7161	TCCAGAAATTTCCCGCGGCAATGCCCATTATGGGAGCGCCGGAATTACAACCTCGACTG	7220
QY	241	CTGAGTCTCTGAGAGCCCGCGGACTACGTCCTCCAGTGTGACATGAGTGGCCACTGCGCA	300
Db	7221	ATAAGTCTCTGAGAGGACCGCGGACTACGTCCTCCAGTGTGACAGGAGTGGCCACTGCGCA	7280
QY	301	CCCTACTAAGACCCCTCCTATACACCTCCACGGAAGAAAGGACAGTTGTTGACAGAA	360
Db	7281	CTTACCAAGGCCCCCTCAATACCTCCACGGAAGAAAGGAGTGTCTTGACAGAA	7340
QY	361	TCCACGCTCTCTTGCCTCTGCGGAGTGTGCACAAAGCTTTTGTAGTCCGACCG	420
Db	7341	TCCACGCTCTCTTGCCTCTGCGGAGTGTGCACAAAGCTTTGAGTCCGACCGATG	7400
QY	421	TGGCGCTGACAGCGGACAGCGAACCGCCCTCTGACCAATCTCTCGACAGCGCGGA	480
Db	7401	TGGCGCGCGAAGAGCGGACAGCGCCCTCTGACCAACCGCCCTCGACAGCGCGGAC	7460
QY	481	GGAAGATCGAGGTGAGTGCATCTCTCAAGCCGCCCTTGAGGGGAGCGGGGGAG	540
Db	7461	GGAAGATCGAGGTGAGTGCATCTCTCAAGCCGCCCTTGAGGGGAGCGGGGGAG	7520
QY	541	CCCGATCTGAGGAGCGGATCTTGATCTACCGTGAAGAGAGAGCGGATGAGGACGTCTG	600
Db	7521	CCCGATCTGAGGAGCGGATCTTGATCTACCGTGAAGAGAGAGCGGATGAGGATGCTG	7580
QY	601	TGCTGCTGATGCTCTTACATGAGACAGGCGCTTGATACGCGCATGCGCTCGGAGAA	660
Db	7581	TGCTGCTGATGCTCTTACATGAGACAGGCGCTTGATACCAATGCGCTCGGAGAA	7640
QY	661	AGCAAGCTGCCATGAAGCGGTGAGCAACTTTGCTGTCGTCACCAACAATGCTAC	720

Db 7641 AGCAAGTATCCCATCATGCTGTTGAGCAACTTTGTGCGCCACCATTAATGCTAT 7700  
7721 GTTACCAATCCCGGAGGCGGAGGCGGAGGAGAGTACCTTTGACAGCTGCA 780  
7701 GCGACAAATCCCGGAGGCGGAGGCGGAGGAGAGTACCTTTGACAGCTGCA 7760  
7781 ATCTGAGAGATCACTACGAGAGCTGTCTCAAGAGATGAAGGCGGAGGCTTCA 840  
7761 GTCCTGAGAGCACTACGAGAGCTGTCTCAAGAGATGAAGGCGGAGGCTTCA 7820  
841 AAGGCTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACCCGCCATTTGCGCA 900  
7821 AAGGCGAGAGCTCTATCCGTAGAGAGGCGCTTAAGCTGACCCGCCATTTGCGC 7880  
901 TCTAATTTGGCTATGAGGCGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAG 960  
7881 TCCAAATTTGGCTATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7940  
961 ATCCGCTCCGTGTGAGAGAGCTTGTGAGAGACACTGAACCAATTTGACACCAATC 1020  
7941 ATCCGCTCCGTGTGAGAGAGCTTGTGAGAGACACTGAACCAATTTGACACCAATC 8000  
1021 ATGCGAAAAATGAGGTTTTCGCTCCACAGAGAGAGAGAGGCGGAGGAGGAGGAG 1080  
8001 ATGCGAAAAATGAGGTTTTCGCTCCACAGAGAGAGAGAGGCGGAGGAGGAGGAGGAG 8060  
1081 CTATCGTGTCCAGAGCTTGGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
8061 CTATCGTGTCCAGAGCTTGGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8120  
1141 GTCTCACCTCCCTCAGGCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
8121 GTCTCACCTCCCTCAGGCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8180  
1201 CAGCGGCTGAGATTCCTGCTGAAAGCTGGAATCAAGAGAGAGGAGGAGGAGGAGGAG 1260  
8181 CAGCGGCTGAGATTCCTGCTGAAAGCTGGAATCAAGAGAGAGGAGGAGGAGGAGGAG 8240  
1261 TATGACACCGGCTGTTGAGACTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
8241 TATGACACCGGCTGTTGAGACTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8300  
1321 ATTATCAATGTTGAGACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
8301 ATTATCAATGTTGAGACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8360  
1381 CGGCTTTATATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
8361 CGGCTTTATATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8420  
1441 TSCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500  
8421 TSCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8480  
1501 GCGCTTGAAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
8481 GCGCTTGAAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8540  
1561 GCGCTTGTGCTTATCTGTGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
8541 GACCTTGTGCTTATCTGTGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8600  
1621 TTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680  
8601 TTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8660  
1681 GACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740  
8661 GACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8720  
1741 AAAAGGATTAATCACTACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1769

Db 8721 AAAAGGATTAATCACTACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8749

RESULT 11  
AB049091  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCES  
AUTHORS  
TITLE  
JOURNAL

AB049091 9547 bp RNA linear VRL 22-AUG-2002  
Hepatitis C virus gene for polyprotein, complete cds,  
isolate:HCV142.  
AB049091  
AB049091.1 GI:11559448  
Hepatitis C virus  
Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.  
1  
Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K.,  
Hatahara, T., Ohita, Y., Kanai, K., Maruo, H., Baba, K., Hijioka, M. and  
Mishiro, S.  
Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients  
with hepatocellular carcinoma: the 'progression score' revisited  
Hepatology. Res. 20 (2), 161-171 (2001)  
11348851  
2 (bases 1 to 9547)  
Mishiro, S.  
Direct Submission  
Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,  
Department of Medical Sciences; 6-3-22 Higashi Oh-i, Shinagawa-ku,  
Tokyo 140-8532, Japan (E-mail: shunji.mishiro@po.toshiba.co.jp,  
Tel: 81-3-3764-8981, Fax: 81-3-3764-8992)  
location/Qualifiers  
1. .9547  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/isolate="HCV142"  
/specific\_host="human"  
/db\_xref="taxon:11103"  
/tissue\_type="serum"  
/note="HCV-genotype: 1b"  
1. .259  
/note="partial"  
260. .9292  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="BA18804.1"  
/db\_xref="GI:11559449"  
/translation="MSTNPKRQRTKXNTRRPODYKFGGQGVGVYLLPPRGPRL  
GVATYRTTSRSGRRQPTPKARQPEGRMAQPGYWPVLGNEGLGMAWGLSPRG  
SRPSPGPTDPRRRRLNGKVIDLTGCPADLMGYIPLVGAPLGGAALAGVLEAD  
GVNATGNLPGCSRSIPLALISCLTIPAAVEYRANSGIYHVTNDCSNSIYVEAD  
VIMHAPGCVPCVRNNSRCVVALPTLILARNSEVPTTIRRHVDLVTGAFCGAAV  
VGDIQGVPLISOLFSPSRHETVQSCSIVGHGRMAMMNMNSPTALIV  
SOLIRPQAVMDVYAGAHGVLALAYYSNMAKYLIVMLPAGVDGTHVTGAGQ  
GSRAPRLSPFTSGASQKQIQLINTSGMHNIRLNLNCDLSHTFLALFTHFNAS  
CEPERLASCRSIDFDQMGPIYAOSSSDORPCWYAPPCGIYPAVQVPCPVYC  
ETPSPVVGITDRGEPYTGMEIETVLILNTRPQGMWFGCTMNSGFTCCG  
PCNIGGAGNNTLFCPTDCFRKPEAEATYTCGSPMLTPROMDVPRLMILPCTG  
TIPKRVYGVGVHRLAACNMTRGRCNLIEDPRSLSPULLSTTEWMLIPGSEFTLL  
PALSTGLIHONVVDVYLYGVSAVSVLYIKREYLLPLILADRVCAQIMMLL  
IAQAELEMLVILNLSVAGAHGISLFLVFCAMYIKGLVGAALVAYGVGVLPL  
LILALPRAVAMDREMASSCGGVFVGLITLSPYKFLAKLIMVLYEFTIABEH  
LOVWPPLNVRGGDAIILTCALHPELIDITGLALIFGLMVLDAQTRVYFVR  
AAGLIRACMLVRKAVGGHYOMAFMKLALIGTVVDHLIRPMARHGLRDLAAVE  
PVYSDMETKITITGATPAACGDISGLPVASRAGREIILGAPDSLGGGRLAIPIT  
AYSQITRGLACITITSLTGDKNOVBESGVVSTATSFLATCNVCMTYHAGSK  
TLAPKSPITQMYTNVDQDLVGMQAPGASVLPCCTGSSDLVLYTHADVIYRRRG  
DSRGLSLSPRPVSLKSSGGLPCPGVWGLFRAAVCTRGVAKAVDPAVEVMEET  
MRSPVFTDNSPSPAVPOTFOVAHLHAPFGSKSVPAAYAAQYKYLVLNPSVAATL  
GFGAVMSKAGTDNIRGRLITITGAPITYSTYGVFLADGGCGGAYDIIICDCHS  
TDSYITIGIGIVLDOAETAGARLVLTATPRGSLVYVPHMIEVALSNTECIPEYK  
ALIPETAIKGRHILFCHSKKKCBELAKLSLSGINAYITRGLDVSYITPSGAVVVA  
TDALMTGYTDIDSVINDCNCVTQTVJFSLDPTFTITTTVTPDAVSRTOGRGTGCG  
RGIYRFTVTRGERSGMDSVILECEYDAGCAWELTPAETSVRLRAYLNTPGIPVCO

DHLEFVSVFTGLTHIDAHPLSOTKQAGNPPYLVAYOATGACBAQAPPSMDOMK  
LIRKPTLHGPPIILYRLGAVONENVTLTPVTKYIMACASADLEVTTWVUGTILA  
ALANCTLTGTSVIVYGRITLISGPAVTPREVALYREFDEMEBCASHLPIIEGMOIAB  
QKQKALGLQATKQAEAAAVESKPMALFAMKHMNFIISGQYLAGJSTLPGN  
PAIASIMAFATASPLTQSTLFINILGAWAQAAPSAASAFGAGIAGAAVSI  
GAGKVLVDILAGAGVAGALVAFKVMSGMPSTEDLVNLPALISPGALVGVVCA  
ILRBHVGPBGAVOMNRLIAPASRGNHVSPTHVYESDAARVQIISLITOLK  
RLHOMINDECSPPSGSMRLRWMDICTYLIDPKTWLSKILPRLDGVPEFCORGY  
GVWRGDIMHTTGPCGAQJAHVKNKSNMIVGKTSNTHGTFPINATYTPCTPSP  
ANYSRALMRAVAEEVEYTRVDYHITGEMTDNICKCQVPAPEFTFVEVGRILR  
YAPACKPLRBEAVFQVGNLYVSGOLCEBPDAVALVITPDSHATATKATK  
AGSPMSASASOLASPLKAACTTHNSPDLIELMLMRONGANTHREVEN  
KVIILSDPLRAEDEREVSYPAILRHSRNPMPMPWADVPMLPESKMDY  
VPPVHGCPLPTKAPRIPPRKRVULTESTVSALAEIATKTHGSESSAADSCT  
ATAPDQASDDGKESDVEISMPLEBPGDPLSDSMTSVSEBATEADVCCMS  
YTMGTALITPCAEBSKLPINALNSLILHHNNVATTSRSMILSKKTKPRLQVLD  
DHYRDVLKEMKAKASTVKAKLISVEBAKCTLPKHSKAFEGAGAKRMLISKAJHI  
RSMKOLDEDTETPIDITIMAKNEVCVOPEKGRPALIIVPDGLGVSEKMAIYD  
VYSTLPQAVMSSYGFQYSGRQVREVLVYTWKSKKPCNGYRCRASPGLTSCGNTL  
TESIYOCCLABEABQAIRSLRRLYIGSPMNSKONGCYRCPASGVLTSCGNTL  
TCLKATACRAKADQCTMLVNGDLVVICBSAGTQEDPANIRVETEMATRSAPRG  
DPPQPEYDELITSCSSNVSAHADASKRVIYITRDPTPLSRAMETARHPVNSML  
GNIMVAPTLAMAMILMTHFSLILAOBELKALDCQIAGACSTIEPLDPIIIEBLH  
GLSAFSLHSYSPERINRVASCLKGLGVPLRVHRBRASVRAKLTISOGRAATCGKYL  
FMMVAVTKLKLPIPAASQDLDSGMFVAGSGDIYHLSBRAPRPMCLILLSGV  
GTYLBNR"

3'UTR

9293..9331

/note="upstream of poly-U stretch"

3'UTR

9332..9425

/note="poly(U) length is indistinct

3'UTR

9426..9547

/note="3' X tail"

ORIGIN

Query Match 88.3%; Score 1562.6; DB 14; Length 9547;

Best Local Similarity 92.7%; Pred. No. 0;

Matches 1640; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

1 CAAATGACTTCCAGACGCTGACCTCATCGAGGCCAACCTCTGTGGCGGCATGAGATG 60  
Db CATCATCAGTCCCGGAGCCTTGACCTCATCGAGGCCAACCTCTGTGGCGGCAGAGATG 6970  
Qy GGGGGGGAATTAACCGCGTGAAGTCAAGAACAGAGTAACTCTGGAATCTTTTGAC 120  
Db GGGGGGGAATTAACCGCGTGAAGTCAAGAACAGAGTAACTCTGGAATCTTTTGAC 7030  
Qy CCGCTCGAGCGGAGGAGATGAGCGGGAAGTCCGTCGCGGCGAGATCTCGGCGGAAA 180  
Db CCGCTCGAGCGGAGGAGATGAGCGGGAAGTCCGTCGCGGCGAGATCTCGGCGGAAA 7090  
Qy TCCAGAAATATCCACAGAGATCCCGCATGGGCAAGCCGCGATTACAACCTCCGCTG 240  
Db TCCAGAAATATATCCCGAGAGATCCCATATGGGCAAGCCGCGATTACAACCTCCAGTG 7150  
Qy CTGGAATCTTGAAGGCGCGGACTACGTCCTCCAGTGTACATGGTGTCCCATGCA 300  
Db CTGGAATCTTGAAGGCGCGGACTACGTCCTCCAGTGTACATGGTGTCCCATGCA 7210  
Qy CCTACTAAGACCCCTCTATACACTCAACGGAAGAAAGAGACAGTGTCTGACAGAA 360  
Db CCTACTAAGACCCCTCTATACACTCAACGGAAGAAAGAGAGTGTCTGACAGAG 7270  
Qy TCCACAGTGTCTTCTGCTCGGCGAGCTTGCCACAAAGCTTTGTAGTCCGACCG 420  
Db TCCACAGTGTCTTCTGCTCGGCGAGCTTGCCACAAAGCTTTGTAGTCCGACCG 7330  
Qy TGGGCGTGAAGCGGCAAGCGCAACCGCCCTCTGACCAATCTCTCGACGAGCGGGA 480  
Db TGGGCGTGAAGCGGCAAGCGCAACCGCCCTCTGACCAAGCTCTGACGAGCGGCGAC 7390  
Qy GAGAGTCTGAGCTTGTATCTCTCATGCCCCCTTTGAGGGGAGCGCGGGGAG 540  
Db GAGAGTCTGAGCTTGTATCTCTCATGCCCCCTTTGAGGGGAGCGCGGGGAG 8530

7391 AAGAAATCCGAGCTTGAAGTGTACTCTCATGCCCCCTTTGAGGAGAGCGGCGGAG 7450  
Qy CCGGATCTAGAGAGCGGCTTGTGTCTACCGTGAAGTGAAGGCGGTGAAGACGTCTGC 600  
Db CCGGATCTAGAGAGCGGCTTGTGTGTCTACCGTGAAGTGAAGGCGGTGAAGACGTCTGC 7510  
Qy TGGTGTGATGCTCTACATGAGACGAGGCGCTGTATGACCGCATGCGCTCGAGAGAA 660  
Db TGGTGTGATGCTCTACATGAGACGAGGCGCTGTATGACCGCATGCGCTCGAGAGAA 7570  
Qy AGCAAGCTCCCATCAACCGCTTGAGCAACTTTGCTGCGTACCAACAATGTCTAC 720  
Db AGCAAGCTCCCATCAACCGCTTGAGCAACTTTGCTGCGTACCAACAATGTGTAT 7630  
Qy GCTACCAATCCCGAGCGCAAGCCAGCGGAGAAAGGTCACTTTGACACATCTGCA 780  
Db GCTACCAATCCCGAGCGCAAGCCAGCGGAGAAAGGTCACTTTGACACATCTGCA 7690  
Qy GGCACAACTCCCGAGCGCAAGCCAGCGGAGAAAGGTCACTTTGACACATCTGCA 7840  
Db GGCACAACTCCCGAGCGCAAGCCAGCGGAGAAAGGTCACTTTGACACATCTGCA 7750  
Qy AAGGCTAAGCTTCTATCACTAGTAGAGAGGCTGCAAGCTGACGCCCCCATTCGGCCAA 900  
Db AAGGCTAAGCTTCTATCTGTAGAGAGGCTGCAAGCTGACGCCCCCATTCGGCCAA 7810  
Qy TCTAAATTTGGCTATGAGGAGAAAGAGAGCTGCGGAACCTTCCAGCAAGCCATTAA 960  
Db TCTAAATTTGGCTATGAGGAGAAAGAGAGCTGCGGAACCTTCCAGCAAGCCATTAA 7870  
Qy ATCCGCTCGGTGAGGAGAGCTTGTGAGAGACCTGGAACAACAATTGACACCAACAC 1020  
Db ATCCGCTCGGTGAGGAGAGCTTGTGAGAGACCTGGAACAACAATTGACACCAACAC 7930  
Qy ATTCGCTCGGTGAGGAGAGCTTGTGAGAGACCTGGAACAACAATTGACACCAACAC 7930  
Db ATTCGCTCGGTGAGGAGAGCTTGTGAGAGACCTGGAACAACAATTGACACCAACAC 7930  
Qy ATGCAAAATAGAGTTTCTGCGTCAACAGAGAGAGAGGCGGCAAGCGAGCTCGC 1080  
Db ATGCAAAATAGAGTTTCTGCGTCAACAGAGAGAGAGGCGGCAAGCGAGCTCGC 7990  
Qy ATGCAAAATAGAGTTTCTGCGTCAACAGAGAGAGAGGCGGCAAGCGAGCTCGC 7990  
Db ATGCAAAATAGAGTTTCTGCGTCAACAGAGAGAGAGGCGGCAAGCGAGCTCGC 7990  
Qy CTATGCTTCCAGACTTGGGAGAGTCCGTATGCGAGAAATGAGCCCTTTATGACGTA 8050  
Db CTATGCTTCCAGACTTGGGAGAGTCCGTATGCGAGAAATGAGCCCTTTATGACGTA 8050  
Qy GTCTCAGCCTCCCTCAGGCTGAGTGGGCTCCTGTAGAGATTCAGATTCCTGGA 1200  
Db GTCTCAGCCTCCCTCAGGCTGAGTGGGCTCCTGTAGAGATTCAGATTCCTGGA 8110  
Qy CAGCGGCTGAGATCTCTGCTGAACGCTGGAATCAAGAAACCCCTATGAGCTTTGCA 1260  
Db CAGCGGCTGAGATCTCTGCTGAACGCTGGAATCAAGAAACCCCTATGAGCTTTGCA 8170  
Qy TATGACACCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 1320  
Db TATGACACCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 8230  
Qy ATTATCAATGTTTGAATTTGGCCCCCGAAGCAAGCAAGCCATAAGTGTGCTCAAGAG 1380  
Db ATTATCAATGTTTGAATTTGGCCCCCGAAGCAAGCAAGCCATAAGTGTGCTCAAGAG 8290  
Qy CCGCTTATATCGGGGCTCCCTGACTAATTAAGAAAGGCGAGAACTGCGGCTATCGCCG 1440  
Db CCGCTTATATCGGGGCTCCCTGACTAATTAAGAAAGGCGAGAACTGCGGCTATCGCCG 8350  
Qy TGCAGGAGAGGCGCTGCTGACAGTACGCTGCGGTAATACCTTCACATGTTA 1500  
Db TGCAGGAGAGGCGCTGCTGACAGTACGCTGCGGTAATACCTTCACATGTTA 8410  
Qy GCTCTGAGAGCTGCTGAGAGCTGCAAGAGCTGCAAGAGTGTGCTGAGAGAG 1560  
Db GCTCTGAGAGCTGCTGAGAGCTGCAAGAGCTGCAAGAGTGTGCTGAGAGAG 8470  
Qy GGCCTGTGCTTATCTGAGAGAGCGGAGAAACCAAGAGAGAGCGGCGAGCTTACAGATC 1620  
Db GGCCTGTGCTTATCTGAGAGAGCGGAGAAACCAAGAGAGAGCGGCGAGCTTACAGATC 8530

QY	1621	TTCAAGGAGCTATGACTAGTACTCTGCCCCCCCCGGGAGACCCGCCCAACAGATAC	1680
Db	8531	TTCAAGGAGCTATGACTAGTACTCTGCCCCCCCCGGGAGACCCGCCCAACAGATAC	8590
QY	1681	GACCTGAGCTGATTAACATCATGCTCTCTCAATGTGTCTGCGCGACATCATCTGGC	1740
Db	8591	GACCTGAGCTGATTAACATCATGCTCTCTCAATGTGTCTGCGCGACATCATCTGGC	8650
QY	1741	AAAAGGCTACTACCTCACCGGTACCC	1769
Db	8651	AAAAGGCTACTACCTCACCGGTACCC	8679
RESULT 12			
AF207753			
LOCUS	AF207753	9374 bp	RNA linear
DEFINITION	Hepatitis C virus strain MD12 complete genome.		VR1 27-APR-2000
ACCESSION	AF207753		
VERSION	AF207753.1	GI:7650223	
KEYWORDS			
SOURCE			
ORGANISM	Hepatitis C virus		
	Hepatitis C virus		
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
REFERENCE	1 (bases 1 to 9374)		
AUTHORS	Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Marumo, F. and Sato, C.		
TITLE	Characteristics of hepatitis C viral genome associated with disease progression		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 9374)		
AUTHORS	Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and Sato, C.		
TITLE	Direct Substitution		
JOURNAL	Submitted (23-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan		
FEATURES			
source	1..9374	Location/Qualifiers	
	/organism="Hepatitis C virus"		
	/mol_type="genomic RNA"		
	/strain="MD12"		
	/db_xref="taxon:11103"		
	/note="isolated from a patient with persistently-normal ALT values (ASC-2)"		
	330..9362		
CDS			
	/codon_start=1		
	/product="polyprotein"		
	/protein_id="AAF65943.1"		
	/db_xref="GI:7650224"		
	/translation="MSTNPKPQKTKRNTNRPRPOVKKPPGGQIVGVYLLPRRPRLLGVRATKTSERSQPRGRQPIFKRQPSRTAQAQPEYPLVINGEMWAGMLSPRGSRPWPDPRRRRRLKVIDLTCGFADLMGYLPVGAAPGARVLAHGRVLEDMVATAGNLPGCSFSLFIALSLCLTTPASAEVRNVSMTHTVDCSSISYAADGMHAAGCVPCRENNSSRCVVALPTLALRNASVPTTIRRHVLLVCAAFSCAMYVDDLSGVFLISOLFTFSPRHRHTVDDNCSTIPGHSRHRAMMMNMWSTLSTVSGLRIPQAVVAVAGAHGWVLAGLAAYISAGNNAKVLVLLPAGVNGQRTISGTOGHTTSGLVLFSSGSQNTQLINTNGSMINTALNCNDLSTGLFALVYTHKNASGCPARNAICRSIDTEPAQGWGPTVTAQPSIDSRPYCMHAAPQCGIIVASOVCGGVCFPPSPVAVGTTDRSDRGVPTVSGENETDVLINLNRBPPOGNMGCTMMSTGCTKCGPCPNIGVGNNTLTCPTDCFRKHPKATYKCGSGPMLFRGLVDVPLRYMHYPCVNFSTFKRVYGVGVEHRLNAAQNMTRGRCGLEDRDSLESLPLLSTTBQVLPVPCSTTLLPALSTGLIHLHONVDVQVLYIGISAVVSPFYKMDYVYLPLILDARCACTMMMLLTAQBALENLVLANASVAGAGPLSLVFCAMVYKGRVLPGAALFYHFMVLLLLALPRAVYAMDREMAASCGAVLVGLMLTSLPYKVFARLIMVLOYLITRAEHLGVWVPLNVRGRDAIIFLTCAHPELFDITKLLALILGLMVLQIGITFVYFVARAGLIRACMLVRKAGHYVOMAEMKALATGTYYDHLTILPGNAHGLRDLAAVAVPVSFDMETKIIITMGADTAACGDIILGLPVSARGRRELLIGPADSLBEGMARLLAPITAYSOOTRGLGCLITSUTGRDKQVGEGRVQVSTATQTLATCNAGVCTTGHGASKTTAGPKGPIITQMTNTVDQLVGMQAPGARSWPTCCGSSDLYLTRADVLPVRRGDRGSLPRPVSYLKSSGSGPLCPSGHVGIFRAAVCYRAAVVAVFVPVSEMTTMSPVFTDMSBPAPVQPTFOVAHLHAPTSKSTKVPAPVAAQYKXVLVNLNSVATL		

ORIGIN

GFAYMSKAVGTDPNIRGTGATTTTGATITVSTYKFLADGCGGAVDIIICNECHS  
TDSYTIILGIGVLVLOAETAGARLVLAATATPESVATVPHPIEBEVALSNTEIPFYK  
AIPIETIKGSHLIFCHSKKKDELAALKSLGAINAVAYRGDVSIVLVGGVVA  
TDLMTGTGDPFDSIDONTCTVOTVDPTEPTITPTVPOAVSARTORGRTGG  
RRGIYRTPTEBPSGWFDSVLECDYDAGAMVELTPAETSVRTBAYITRPGVPCO  
DHLFMEBVFGLTRIDHPLSTQTKQDNPIVAVIQAIVCARQAAPPSMDMKC  
LTRIKPLTHGPTPLRYLGAOVNVTLLHPTKXIMTMSADLEIVSTVWLVGVA  
ALAAVCLITGSSVIVGR11LISGRBNIIPDEVLREBDEMECSQIPLYEQMLAB  
OFKQALQIOTATKOAAAPVYESKRAALAEFMWGMNFIISGIQYLAGSLTLPN  
PAISLMAFTASVTSPLTOSTLTFNLIGVNAOLAIPSAASFVAGAGIAGAVSI  
GLGVYVDIILAGYAGVAGALVAKMSGMBSTEDVNLILPALSGLAVYGVCA  
ILKRVHGBGAVQVMMNRLLIFASRGHVSFTHVPSDDAARVTOQLNSLITTOILR  
RLHOMINEDSCPCSGSMLRDVMDICTVLSDFETWLSKILPLPGVPLSCRGYK  
GVNRGDIGIMOTTCPGCAOITGHVKNQSMRIVGPTCSTNTHGTPPINAYTGPCTSP  
APNSRBLMRYAAEEYVEVTEGPHVYVGTNTDNCPCGCPAPBEFTEDVGRHLR  
YAPACKPLREBYFOVGLNOXYVGSOLPCGPEPDVAVLSMLTDPHITETAKRL  
AKSGPPSLHSSASQLSAPSLKATCTTHKSPDDDLBAILMRQENGNITRRESR  
KVLILDSFDPLRAEDEREISVPAELIRKPKFPALPIMARPDYNPILSWMKPDY  
VPVPHGCLPPTKGPPLPPRRRRTVTLTESTVSLAEATTFGSSSGADSGT  
ATAPDDQSDGSDVSEYSMPLEGEPRDLDLSGMSVSEAGBEDVCCMS  
YTMGALITPCAASESKLPIINALNSLLRHNNMYATTSRACOROKKVFDRLOYLD  
DHYRDLVKEMKAKASTYKAKLSVEBAKCTTPHSASKRGYAKQVNRNLSKRVNH  
RSWKDLIEDTEPTINTTWAKEVFCYQPKGRKARLIVPDLGVCEKALVD  
VSTLPQAVWSSSYFOYSPQORVEPLVNTWKSCKCPGSEYDTRCDSVTENDIRV  
BESLYOCDLAPEAROVIRSLTERLYVGGPLTNKGQNCGRRCRAGSVLTSGNTL  
TCYLKASACRAALADCTMLVNDGDLVLCESAGTOEDALSLVFTEAMRYASAPG  
DPEOEXDELETSCSSNVSAVHADSGRYYLFRDPTPLAPARAMEATRPVNSML  
GNITMVAPTLMABMILMTHPSTILLAOELEKALDQIYGACVSIIEPLDIPQIIEHL  
GLSAPLSHSYSPGEINRYASCLRLGCTPPLRVHHRARSRAKLSQGRATVSGKL  
FMAVRITKLPTIPASRLDLSGMFVAGISGDVIYHLSRARPRMCLLTLISVG  
GYILLPNR"

Query Match 88.0%; Score 1557.8; DB 14; Length 9374;

Best Local Similarity 92.5%; Pred. No. 0; Mismatches 132; Indels 0; Gaps 0;

Matches 1637; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY	1	CMAATGACTCCAGAGCGTGAAGCTCATCAAGGCAACCTCTGTGCGGAGATG	60
Db	6981	CACATGACTCCCGGAGCGTGAAGCTCATCAAGGCAACCTCTGTGCGGAGATG	7040
QY	61	GCGGGGACATTACCGCGTGAAGTCAGAGAACAGGATGTAATCTTGACTCTTTCGAC	120
Db	7041	GCGGGGACATTACCGCGTGAAGTCAGAGAACAGGATGTAATCTTGACTCTTTCGAT	7100
QY	121	CCGCTCCGAGGAG	180
Db	7101	CCGCTCCGAGGAG	7160
QY	181	TCCAAGAAATTCACACAGCATGCCGATGAGGACACCGCGGATTAACCTCTCCGCTG	240
Db	7161	CCGAGAGATTCCTCCCGAGCTTCCCATATAGGACACCGCGGATTAACCTCTCCATG	7220
QY	241	CTGAGTCTTGAAAGGCGCGGACTACGTCCTCTCAAGTGTAATCGAGTCCCACTGCA	300
Db	7221	CTGAGTCTTGAAAGGCGCGGACTACGTCCTCTCAAGTGTAATCGAGTCCCACTGCA	7280
QY	301	CCTACTAGACCCCTCTCTATACCACTTCCAGAGGAAAGAGACAGTTGTTCTGACGAA	360
Db	7281	CCTACTAGACCCCTCTCTATACCACTTCCAGAGGAAAGAGACAGTTGTTCTGACGAG	7340
QY	361	TCACCGGTCTTCTGCGCGGAGAGCTTCCCAAAAGCTTTGGATGCTCCGAGACG	420
Db	7341	TCACCGGTCTTCTGCGCGGAGAGCTTCTCTAAGACCTTTGGACAGCTCCGGGTG	7400
QY	421	TCGGCGTGCAGACGCGGACGCGCAACCGCCCTCTCTGACCAATCTCTGACGACGCG	480
Db	7401	TCGGCGCGCGACAGTGGACGCGGACCGCCCTCTGATGACACTTGAAGACGCGGAC	7460
QY	481	GCAAGATCTGACGTTGATCTGATCTCTCATGCCCCCTTGAAGGGGACCGGGGAC	540
Db	7461	AAAGGATCGACGTTGATCTGATCTCTCATGCCCCCTTGAAGGAGACCGGGGAC	7520



```

QY 541 CCCGATCTCAGCAGCGGGTCTTGCTCTACCGTGAAGGAGGCGGGTGAAGCGTCTC 600
DB 7521 CCGGATCTCAGCAGCGGGTCTTGCTCTACCGTGAAGGAGGAGGCGGGTGAAGCGTCTC 7580
QY 601 TGTGCTGATGTCCTACACATGAGAGGCGCTCTGATCAAGCAGCGCTGCGAGGAA 660
DB 7581 TGTGCTGATGTCCTACACATGAGAGGCGCTCTGATCAAGCAGCGCTGCGAGGAA 7640
QY 661 AGCAAGCTGCCATCAACCGCTTGAGCAACTCTTGTGCTGCTACCAACATGCTTAC 720
DB 7641 AGCAAGCTGCCATCAACCGCTTGAGCAACTCTTGTGCTGCTACCAACATGCTTAC 7700
QY 721 GGTACACATCCCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 780
DB 7701 GGTACACATCCCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 7760
QY 781 ATCTGAGCAGTACTACAGAGAGCGTGTCAAGAGATGAGAGCGAGCGGAGCGGAGG 840
DB 7761 GTCCTGAGAGCAGTACTACAGAGAGCGTGTCAAGAGATGAGAGCGAGCGGAGG 7820
QY 841 AAGCTAAGCTTCTATCAGTAGAGAGAGCGCTGCAAGCTGACGCCCGACATTGCGCA 900
DB 7821 AAGCTAAGCTTCTATCAGTAGAGAGAGCGCTGCAAGCTGACGCCCGACATTGCGCA 7880
QY 901 TCTAATTTGGCTATGAGGAGAGAGCGTCCGAGCTATCCAGAGCGCATTAACAC 960
DB 7881 TCTAATTTGGCTATGAGGAGAGAGCGTCCGAGCTATCCAGAGCGCATTAACAC 7940
QY 961 ATCCGCTCCGTGTGAGAGAGACTTGTGAGAGACACTGAAACAAATTGACACACACATC 1020
DB 7941 ATCCGCTCCGTGTGAGAGAGACTTGTGAGAGACACTGAGACACATTAACACACATC 8000
QY 1021 ATGGCAAAATAGAGTTTCTGCTGCAACCAAGAGAGAGCGGAGCGGAGCGGAGCTCGC 1080
DB 8001 ATGGCAAAATAGAGTTTCTGCTGCAACCAAGAGAGAGCGGAGCGGAGCGGAGCTCGC 8060
QY 1081 CTATGCTGTCCAGACTTGGGGGTCCGCTGTGAGAGAAATGGCCCTATGAGCGT 1140
DB 8061 CTATGCTGTCCAGACTTGGGGGTCCGCTGTGAGAGAAATGGCCCTTATGAGCGT 8120
QY 1141 GTCTCAGCCTCCCTCAGGCTGATGAGGCTCTCTGATCAGATTCAGATTCCTCTGCA 1200
DB 8121 GTTTCACCTTCTCAGGCGGTGATGAGGCTCTCTGATCAGATTCAGATTCCTCTGCA 8180
QY 1201 CAGCGGCTGAGTCTCTGTGAAAGCGCTGAAATCAAGAGAGCGGCTATGCGCTTGA 1260
DB 8181 CAGCGGCTGAGTCTCTGTGAAATCAAGAGAGCGGCTATGCGCTTGA 8240
QY 1261 TATGACACCGGCTTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 1320
DB 8241 TATGACACCGGCTTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 8300
QY 1321 ATTATCAATGTTGATCTTGGCCCGGAGAGCAGACAGGCGATTAAGTGTCTCAGAG 1380
DB 8301 ATTATCAATGTTGATCTTGGCCCGGAGAGCAGTCAAGAGTCAATAGTGTCTCAGAG 8360
QY 1381 CGGCTTATATCGGGGCTCCCTGACTAATTCAAAAGGAGAGATCGCGCTATCGCGG 1440
DB 8361 CGGCTTATATCGGGGCTCCCTGACTAATTCAAAAGGAGAGATCGCGCTATCGCGG 8420
QY 1441 TGGCGGAGAGCGGCTGCTGAGACTAGAGTGGGATATACCTCTCAGATTCCTTGAG 1500
DB 8421 TGGCGGAGAGCGGCTGCTGAGACTAGAGTGGGATATACCTCTCAGATTCCTTGAG 8480
QY 1501 GCGTCTGAGCGCTGAGCTGCAAGCTCAAGAGCTGACAGTGTCTGTGCGAGAG 1560
DB 8481 GCGTCTGAGCGCTGAGCTGCAAGCTCAAGAGCTGACAGTGTCTGTGCGAGAG 8540
QY 1561 GCGCTTGTGTTATCTGTGAGAGCGGAGAACCCAGAGAGAGCGCGGAGCTTACAGATC 1620
DB 8541 GACCTTGTGTTATCTGTGAGAGCGGAGAACCCAGAGAGATGCGGAGCTTACAGATC 8600
QY 1621 TTCAGGAGGCTATGACTAGTACTGCGCCCGGCGGAGAGCGGCGGAGAGATAC 1680

```

```

DB 8601 TTACCGGAGGCTATGACTAGTACTGCGCCCGGAGAGCGGCGGAGAGCGGCGGAGAGATAC 8660
QY 1681 GACCTGAGTGTATTAATCATCATGCTCTCTCAATGTGTGCTGCGCAGCATGATCTGCG 1740
DB 8661 GACCTGAGTGTATTAATCATCATGCTCTCTCAATGTGTGCTGCGCAGCATGATCTGCG 8720
QY 1741 AAAAGGTATTAATCACTTCCCGGTGAGCC 1769
DB 8721 AAGAGGTGTACTACTCACTGAGCCG 8749

RESULT 13
HPCGENANTI
LOCUS HPCGENANTI 9425 bp mRNA linear VRL 14-JUL-1994
DEFINITION Hepatitis C virus Taiwan mRNA, 5' end.
ACCESSION M84754
VERSION M84754.1 GI:329763
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (bases 1 to 9425)
AUTHORS Chen, P. J., Lin, M. H., Tai, K. P., Liu, P. C., Lin, C. J., and Chen, D. S.
TITLE The Taiwanese hepatitis C virus genome: sequence determination and
mapping the 5' termini of viral genomic and antigenomic RNA
JOURNAL Virology 188 (1), 102-113 (1992)
MEDLINE 92230206
PUBMED 1314449
COMMENT Original source text: Hepatitis C virus (strain Taiwan) cDNA to
mRNA.

FEATURES
source location/Qualifiers
1..9425 /organism="Hepatitis C virus"
/mol_type="mRNA"
/db_xref="taxon:11103"

ORIGIN
Query Match 88.0%; Score 1557; DB 14; Length 9425;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1635; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 5 ATGACTTCCCAAGCGTGAAGCTGATGAGGCGCAACTCTGTGGGCGGATGAGATGGCG 64
DB 6997 ATACCCCGGAGAGCGTGAAGCTGATGAGGCGCAACTCTGTGGGCGGAGATGGCG 7056
QY 65 GGGACATTACCGCGGTGAGTGAAGAAAGATGATCCGAGCTCTTGAACCGCG 124
DB 7057 GAAACATACCGCGGTGAGTGAAGAAATGATGATCCGAGCTCTTGAACCGCG 7116
QY 125 TCCGAGCGGAGAGATGAGCGGAAAGTGTCCGTCCCGGAGAGATCTCGGAAATCA 184
DB 7117 TCCGAGCGGAGAGATGAGCGGAAAGTGTCCGTCCCGGAGAGATCTCGGAAATCA 7176
QY 185 AGAATTTCCACAGAGATGCGCGATGGGAGCGCGCGATTAACAACCTCCGCTGCTGG 244
DB 7177 GGAAGTTCCTCAGAGTGTCCCGATGAGAGCGCGCGATTAACAACCTCCATGCTGAG 7236
QY 245 AGTCCGGAAGGCGCGGAGTCAAGTCCCTCAAGTGTATCATGAGTCCGAGTCCAGCTTA 304
DB 7237 AGCTCGGAAGGCGCGGAGTCAAGTCCCTCAAGTGTATCATGAGTCCGAGTCCAGCTTA 7296
QY 305 CTAGAAGCGGCTCTATACCACTTCAAGAGAAAGAGACAGTTGTTCTGACAGATCA 364
DB 7297 TCAAGGCGGCTCAATACCACTTCAAGAGAAAGAGACAGTTGTTCTGACAGATCA 7356
QY 365 CCGTGTCTTGTGCGCGGAGCTTGGCAAAAGCTTTTGTGATCTCCGAGCGTCCG 424
DB 7357 CCGTGTCTTGTGCGCGGAGCTTGGCAAAAGCTTTTGTGATCTCCGAGCGTCCG 7416
QY 425 CCGTCAAGCGGAGCGGAGCGGCGGCGGCGGAGCGGCGGAGCGGAGCGGAGAG 484

```

Db 7417 CCGCTGTAGACGGGACCGGACCGCCCTCTGTATCAGCCCTCCGACGACGGGACGACG 7476  
Oy 485 GATCTAGACGTTGAGTGCATCTCTCATGATGCCCCCTTGAGGGGGAGCGGGGGACCCCG 544  
Db 7477 GATTCGACGTTGAATCTGCTCTCTCTATGCCCCCTTGAGGGGGAGCGGGGGACCCCG 7536  
Oy 545 ATCTCAGCAGCGGGCTTGTGTCTACCGTGAAGTGAAGAGGCGGTGAAGACGTGCTGCT 604  
Db 7537 ATCTCAGCAGCGGGCTTGTGTCTACCGTGAAGTGAAGAGGCGGTGAAGAGCGTCTGCT 7596  
Oy 605 GGTTCAGTCTCTACATGAGACAGGGGCTCTGATCAGCGCATGCGGTGGAGAGAAAGCA 664  
Db 7597 GCTCAGATGCTCTACATGAGACAGGGGCTCTGATCAGCGCATGCGGTGGAGAGAAAGCA 7656  
Oy 665 AGCTGCCCATCAACGCGTTGAGCAACTCTTTGCTGCGTCAACCAACATGCTCTAGCTA 724  
Db 7657 AGCTGCCCATCAATGCGTTGAGCAACTCTTTGCTGCGTCAACCAACATGCTCTAGCTA 7716  
Oy 725 CCACATCCCGACCGGACCGGACCGGACGAGAAAGGTCACTTTGACAGACTGCAAAATCC 784  
Db 7717 CCACATCCCGACCGGACCGGACCGGACGAGAAAGGTCACTTTGACAGACTGCAAGTCC 7776  
Oy 785 TGGAGAGTACCTACACAGAGCGTGTCTCAAGAGATGAAGGCGAAAGCGTCCACAGTTAAG 844  
Db 7777 TGGAGAGTACCTACACAGAGCGTGTCTCAAGAGATGAAGGCGAAAGCGTCCACAGTTAAG 7836  
Oy 845 CTAACTCTCTACAGTGAAGAGCGCTGCAAGCTGACGCGCCCACTTCGCGCAATCTA 904  
Db 7837 CTAACTCTCTACAGTGAAGAGCGCTGCAAGCTGACGCGCCCACTTCGCGCAAGTCTA 7896  
Oy 905 AATTGGCTATGGGGGCAAGAGCGTCCGGAACCTATCCAGAAAGCGCAATTAACAACATCC 964  
Db 7897 AATTGGCTATGGGGGCAAGAGCGTCCGGAACCTATCCGGAAGCGCAATTAACAACATCC 7956  
Oy 965 GCTCCGTGTGGAGAGCTTGTGGAAGACACTGAAACACCAATTGACACCACTATCATG 1024  
Db 7957 GCTCCGTGTGGAGAGCTTGTGGAAGACACTGAAACACCAATTGACACCACTATCATG 8016  
Oy 1025 CAAAAAATGAGGTTTCTGCTGCTCAACGAGAGAGAGGCGGCAAGCGAGCTCGCTTA 1084  
Db 8017 CAAAAAATGAGGTTTCTGCTGCTCAACGAGAGAGAGGCGGCAAGCGAGCTCGCTTA 8076  
Oy 1085 TGTGTGTCCAGACTTGTGGGGGTCCTGTGTGCGAGAAATGCGCTCTATGACGTGCT 1144  
Db 8077 TGTGTGTCCAGACTTGTGGGGGTCCTGTGTGCGAGAAATGCGCTCTATGATGTGCT 8136  
Oy 1145 CCACCTCCCTCAGAGCTGATGAGGGCTCTCTGATCGGATTCAGATTTCTCTGACAGC 1204  
Db 8137 CCACCTCCCTCAGAGCTGATGAGGGCTCTCTGATCGGATTCAGATTTCTCTGACAGC 8196  
Oy 1205 GGGTCAGTCTCTGAGTGAAGCGCTGGAATCAAGAAAGACCCCTATGGGCTTTGCAATG 1264  
Db 8197 GGGTCAGTCTCTGAGTGAAGCGCTGGAATCAAGAAAGCGCTCTATGGGCTTTCTCATG 8256  
Oy 1265 ACACCCGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTGAGAGAGTCAATT 1324  
Db 8257 ACACCCGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTGAGAGAGTCAATT 8316  
Oy 1335 ATCAATGTTGTAATCTTGGCCCCCGAAGCGACAGCGCTTAAGGTCCTCAAGAGCGG 1384  
Db 8317 ATCAATGTTGTAATCTTGGCCCCCGAAGCGACAGCGCTTAAGGTCCTCAAGAGCGG 8376  
Oy 1385 TTTATATCGGGGGTCCCTGTAATTAATCAAAAGGCGAATGCGGCTATCCGCGGTGC 1444  
Db 8377 TTTATATCGGGGGTCCCTGTAATTAATCAAAAGGCGAATGCGGCTATCCGCGGTGC 8436  
Oy 1445 GGGCAGCGGGCTGTGACGACTGCGGTAAATACCTTCAATGTTACTTTGAAGGCT 1504  
Db 8437 GGGCAGCGGGCTGTGACGACTGCGGTAAATACCTTCAATGTTACTTTGAAGGCT 8496  
Oy 1505 CTGCGAGCTGTGAGTGCAGAAAGCTCAGAGCTGCAACATGCTGCTGTCGAGAGCGCC 1564  
Db 8497 CTGCGAGCTGTGAGTGCAGAAAGCTCAGAGCTGCAACATGCTGCTGTCGAGAGCGAC 8556

Oy 1565 TTGTGTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACAGAGTCTCA 1624  
Db 8557 TTGTGTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACAGAGTCTCA 8616  
Oy 1625 CGAGGCTATGATGATGAGTCTGTGCCCCCGCGGAGACCGGCCCAACAGAAATGAGC 1684  
Db 8617 CGAGGCTATGATGATGAGTCTGTGCCCCCGCGGAGACCTGCCCAACAGAAATGAGC 8676  
Oy 1685 TGGAGTTGATGATGATGAGTCTGTGCCCCCGCGGAGACCGGCCCAACAGAAATGAGC 1744  
Db 8677 TGGAGTTGATGATGATGAGTCTGTGCCCCCGCGGAGACCTGCCCAACAGAAATGAGC 8736  
Oy 1745 GGGTATCTACCTCAGCCCGTACCC 1769  
Db 8737 GGGTATCTACCTCAGCCCGTACCC 8761

RESULT 14  
AB049090  
LOCUS  
DEFINITION  
Hepatitis C virus gene for polypeptide, complete cds,  
AB049090  
VERSION  
AB049090.1 GI:11559446  
KEYWORDS  
SOURCE  
ORGANISM  
Hepatitis C virus  
Hepatitis C virus  
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE  
AUTHORS  
Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K., Harahara, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and  
Mishiro, S.  
TITLE  
Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the 'progression score' revisited  
JOURNAL  
Hepatol. Res. 20 (2), 161-171 (2001)  
PUBMED  
11348851  
REFERENCE  
AUTHORS  
Mishiro, S.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital, Department of Medical Sciences, 6-3-22 Higashi Oh-1, Shinagawa-ku, Tokyo 140-8522, Japan (E-mail: shunji.mishiro@po.toshiba.co.jp, Tel: 81-3-3764-8981, Fax: 81-3-3764-8992)

FEATURES  
source  
location/Qualifiers  
1. 9573  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/isolate="HCV140"  
/specific\_host="human"  
/db\_xref="taxon:11103"  
/tissue\_type="serum"  
/note="HCV-genotype: 1b"  
<1..310  
/note="partial"  
311..9343  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="BAB18803.1"  
/db\_xref="gi:11559447"

CDS  
translation="MSTNPKRQRTKENTNRPODVKPGGQIVGGVLLPRGRPL  
GVATRTSRSQSGRROPKROPEGAMAGPGVWPLVNGEGMAGMLSPRG  
SRPWSGPDPRSRNIGKVIDTLGCPADLMGIVPAGVILGGAARALAGVLE  
GVNATGNTLPCRPSTILALSLTLTPASAEVRANSGVYHTNDCNSIYVEAD  
MIHTPGCVPCREDNVSRCWALPTLAANSIPLTTIRHVDLVGAALCSAMT  
VGDICGSFVLSQLETPSPRYETVOTCNCSIYGHVSGHMAWMDMMNPPTALVY  
SOLIRIQAIVVDVAGAHMGLVLAAYSMGNMAKYLIVMLFAGVGNHTVGGGA  
GQHARFSLSPSSGASOKIOLINTNGSMHINRTALNCDSLHTGFIALFVAHKENS  
GCPERMASCPIDPGOKGPIISHVNIISQRPYCHVAPRPGGIYAPSGOVPCVY  
FTSPVAVGTTRDFGVFTYTGENTVLLNTRPPOGMFGCTMANGCTFTTCGG  
PCNIGAGANNLTCPDCEPKHPEAVTYKSGSPMLPRCIYDVPYRLMPCVPR  
TTFKRMVVGVEHRLNAACNMTGERCNLEDRSRISLPLSTTEWQIMPSTTL



PAISTGLIHLHONIIVDOYLYGVGSAVSIIVIKMEYVILLFLLADARVACILMMLL  
IAQEAALBENLVYTLNAAISVAGAGTSLTFLVFECAANTIKKIVPGAAIYAFYGTWPLLL  
LLALPRAVAMDRBMAASCGTIVFGLVSLTSLPHTKILALIMLOITFIRACAC  
LOVMIPLINVRGRDAIILLTCAVHPELVFELIKLILALIGPLMVOAGITRYPVVR  
AQQILIRACMLVRVAGGVYVAMLMKLAALATGYVDHLPLMDMAAGELDLVAVE  
PVIFSMETKITIMGADTAAAGDIIILGIPASARGREIILGPDSEGEQMBRLALPT  
AYSGOIRGLIGCIITSLTGRDKMVGEGEVNSTATQSFATQVNCVCMVYHGAASK  
TLAQBKPIITOMVYNDODLUGMPPARGASLTPCQSGSDLYLVRHADVIPVRBC  
DERGSLIFRPISLYLKSGSGPLCPSGHAGVIFRAVCTRGVAKADVPRVSMETT  
MKSVPVINDSSPPAVPQITOVAAIHAPVSGKSTKVPAAAGVYLVLPVAVTL  
FGAVYSKAHGVDPNIRITGRTITTPGATITYSTYKGVPAALVQATVRAQAAPPDMQMKC  
TDSITLIGITVDOAETAGARVATAPRPSGVYVPHNIEVLSNTEIPFGK  
ALPLEAIKGGHILFCHSKKDELAASKGLGVNAVAYYRGDADVIPPSGVVVA  
TQALMGTGDPDSVITDCTCTQYTVDFSLDPFTIETTTPVDPAVSRGORTGRC  
RGTITRFTVPERPSGMLDSVLCECTDQCAATYELTPRPSVRLAALPELPLVQC  
DHLFEVESVFTGLTHIDAFHLSQTKOAGNFPYLVAVQATVRAQAAPPDMQMKC  
LIRKPLETHGPPPLRYLGAQVONEVTLTHITKYIMACMSADLEVTSITVLYGVIA  
ALAAVCLTTGSAVIVGRITLISGKPAIIPDEVLVQEFDEMEBCASHLPYIEQOMLAE  
OFKOKAGLLOTTAKOAAAPVYESKMALEFMAKHMNPIISGIOYLAGSTLLEGN  
PALASIMATFASITSPLOTHTLLPNILGEMVAQOLAPPSASAIPGAGIAGAAVSI  
GAGKVLVDILAGVAGALVAFKIMSGESTEDLVNLLPILISPGALVGVVCA  
LIRHVGPEGAVQVMMNRLIAPASRGHVSPTHYVESDAALVTOILSLITDOLK  
RLHOMINBCSTPCSGSMRLVDWMICTVLVDFKTWLOSKVLRLGVPFLSCORGX  
GVWRGDGIQOTTCPGAOIAGHVNGSMRIVGRTCSNTHGFPINATYTGCTSP  
APVSRALMRVAEEVEITRVGDHNYVTGTTDNIKCPQVAPAPPEPTELDVRLHR  
YAPACKPLRDEYTFOVGNQYVVSQOLPCEPEPDYVTLTSMITDPSHITAEAKRL  
ANGSPSLASSASQSLASBLKATCTTRDOSPADLILEANLRQMGAITRVESEN  
KVILDSFEPRLRAEEDEREVAALILRKSRRPPALPVARWDPIPIESKUDP  
VPPVWGCPLPTKAPPIPPPRKRTIVLESTVSALAEIATKFSGSSAVDSGT  
ATAPPOPSDDGDGADVESYSMSPLLEGPDPLDSDMSVSEASDSSGSCMS  
YMTGALITPCAAESKRLPINLNSNLIHHMNVATRSASROKRYTFLDLOLD  
DHYRDLKEMKAKASTYKAKILSVEACUKLTPHSAKSRKGYAKORYLSAAYNI  
RSWKOLEDOTPIIDTVMAKNEVFCVPEKGRKPARLIVFPDGVRCENALYD  
VUSTIYQAVMGAISYFOYSPGQREVLVNAAMSKRCMPGAVRFDLIVESIDRV  
BESIYOCCLAPARQAIRSLTERLYVGGITNSKQNCYRCRASGYLTSCEGTL  
TCYLKATAACRAKLODCTMLVCGDILVYCESAGOEADLSARVETAMRTASAPG  
DPROPEYDELITSCSNVSVADASGKVVYLTDPDPLAASAEATARHPVNSML  
GNITMYAPTLMAAMILMTHTFSLILAOBOLKALDCOTIGATYSIEPLDPLQIOLH  
GASARSLSHSYSGEINRVASCKLGLVPLRVHRHARSRAULISQGGAACTCGYL  
FMVAVHTKLTLPPIPAASOLDISGMPIAGVSGDIYHLSRARPRFMWCLLHLSGV  
GIYLPNRR"

3'UTR  
9344..9396  
/note="upstream of poly-U stretch"  
9397..9439  
/note="poly(U) length is indistinct  
poly(U) stretch"  
9440..9573  
/note="3' X tail"

3'UTR  
9397..9439  
/note="poly(U) length is indistinct  
poly(U) stretch"  
9440..9573  
/note="3' X tail"

ORIGIN

Query Match 88.0%; Score 1557; DB 14; Length 9573;  
Best Local Similarity 92.6%; Pred. No. 0;  
Matches 1635; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 5 ATGACTCCCAAGCGGTGACCTCATGAGGCGCAACCTCCGTGCGCGGCAATGAGAGGCG 64  
DB 6966 AAGACTCCCAAGCGGTGACCTCATGAGGCGCAACCTCCGTGCGCGGCAATGAGAGGCG 7025  
QY 65 GGGACATTACCGCGTGAAGTCAAGAAACAAGTAGTAATCTGACTCTTTGACCCGCG 124  
DB 7026 GGAACATCAACCGTGTGAGTCAAGAAATAAGTGTGTAATCTTGACTCTTTGAAACCGC 7085  
QY 125 TCCGACCGGAGAGATGAGCGGGAAGTGTCCGTCCCGCGGAGATCTTCGCGGAAATCCA 184  
DB 7086 TTCGGGCGGAGGAGATGAGGGAAGTATCCGTTGCGGCGAGATCTTCGCGGAAATCCA 7145  
QY 185 AGAATTTCCACACGAGGATGCCGCAATGGGCAAGCCCGGATTACAACCTCCGCTGCTG 244  
DB 7146 GGAAGTTCCCCCGGGGCTGCCCCGTATGGGCAAGCCCGGATTACAACCTCCCACTAAG 7205  
QY 245 AGTCTTGAAGGCCCCGGAATGAGTCCCTCAAGTGTATACATGGGTGCCCACTGCACTTA 304  
DB 7206 AGTCTTGAAGGCCCCGGAATGAGTCCCTCCGCTGAGTACACGGGTGCCCACTTGCACCTTA 7265

QY 305 CTAAAGACCCCTCCTATACCACTTCCACGAGAAAGAGACAGTTGTTCTGACAGAAATCCA 364  
DB 7266 CTAAAGACCCCTCCTAAATACCACTTCCACGAGAAAGAGACAGTTGTTCTGACAGAAATCCA 7325  
QY 365 CGGTGCTTCTGCGCTGGCGGAGCTTGGCCAAAGAGCTTTTGGTAGCTCCGAGCCGTGG 424  
DB 7326 CGGTGCTTCTGCGCTGGCGGAGCTTGGCCAAAGAGCTTTTGGTAGCTCCGAGTCCGTGG 7385  
QY 425 CCGTGACACGCGGACGAGGAAACCGCCCTCTTGACCAATCTTCCAGACGAGCGGAGAG 484  
DB 7386 CCGTGACACGCGGACGAGGAAACCGCCCTCTTGATACGCCCTCCATGACGAGAGAGAGCAG 7445  
QY 485 GATCTACGTTAGTGTATCTCTCATGCCCCCTTTGAGGGGAGCCGGGGAGACCCCG 544  
DB 7446 GATCTACGTTAGTGTATCTCTCATGCCCCCTTTGAGGGGAGCCGGGGAGACCCCG 7505  
QY 545 ATCTCAGCAGCGGCTTGTGCTTACCGGTAGTGAAGAGCCGGTGAAGAGTGTGCTGCT 604  
DB 7506 ATCTCAGCAGCGGCTTGTGCTTACTGTAGCGAGAGGCTAGTGAAGATGTCTGCTGTT 7565  
QY 605 GCTCGATGCTTACACATGAGACAGCGCTGTATCAACGCAATGCTGCGAGAGAAAGCA 664  
DB 7566 GCTCGATGCTTACACATGAGACAGCGCGCTGTATCAACCAATGCTGCGAGAGAAAGCA 7625  
QY 665 AGTGGCCCATCAACGCGGTGAGCAATCTTTGTGCTGCTGATCCACCAATGTCTACGCTTA 724  
DB 7626 AGTGGCCCATCAACGCGGTGAGCAATCTTTGTGCTGCTGATCCACCAATGTCTTAAGCA 7685  
QY 725 CCAATCCCGGACGCGGACGAGCGGACAGAAAGAGTCACTTTGACAGACTGCAATATCC 784  
DB 7686 CCAATCCCGGACGCGGACGAGCGGACAGAAAGAGTCACTTTGACAGACTGCAATATCC 7745  
QY 785 TGGACATCACTACAGAGAGTGTCTCAAGAGATGAAGGCAAGCGGTCCACAGTTAAG 844  
DB 7746 TGGACATCACTACAGAGAGTGTCTCAAGAGATGAAGGCAAGCGGTCCACAGTTAAG 7805  
QY 845 CTAAAGCTTATACATGATGAAGAAAGCTGCAAGCTGACGCCCTCACTTGGCCCAATCTTA 904  
DB 7806 CCAAACTTATCTGTATGAAGAAAGCTGCAAGCTGACGCCCTCACTTGGCCCAATCTTA 7865  
QY 905 AATTTGCTATGGGAGCAAGAGAGTCCGGAACCTATCCAGCAAGGCGATTAAACCATATCC 964  
DB 7866 AATTTGCTATGGGAGCAAGAGAGTCCGGAACCTATCCAGCAAGGCGATTAAACCATATCC 7925  
QY 965 GCTCCGTGTGGAGAGACTTGTGGAAGACATGAAACACCAATTGACACCAATCATGAG 1024  
DB 7926 GCTCCGTGTGGAGAGACTTGTGGAAGACATGAAACACCAATTGACACCAATCATGAG 7985  
QY 1025 CAAAAAATGAGGTTTCTGCTGCTCAACAGAGAGAGAGCGGCAAGCCAGCTCGCTTA 1084  
DB 7986 CAAAAAATGAGGTTTCTGCTGCTCAACAGAGAGAGAGCGGCAAGCCAGCTCGCTTA 8045  
QY 1085 TCGTGTCCAGACTTGGGGGTCGCGTGTGGAAGAAATGGCCCTTAATGAGTGTCT 1144  
DB 8046 TCGTGTCCAGACTTGGGGGTCGCGTGTGGAAGAAATGGCCCTTAATGAGTGTCT 8105  
QY 1145 CCAACCTCCCTGAGGCTGTGATGGGCTCTCTGATCGAATTCAGATTCCTCTGAGACAG 1204  
DB 8106 CCAACCTCCCTGAGGCTGTGATGGGCTCTCTGATCGAATTCAGATTCCTCTGAGACAG 8165  
QY 1205 GGGTGAAGTTCTGTGGAAGCGCTGGAATCAAGAAAGACCCCTATAGGCTTTGCATATG 1264  
DB 8166 GGGTGAAGTTCTGTGGAAGCGCTGGAATCAAGAAAGACCCCTATAGGCTTTGCATATG 8225  
QY 1265 ACAACCGCTGTTTGTAGCTGACATGAGATGAGATCCGTGTGAGAGAGTCAATTT 1324  
DB 8226 ACAACCGCTGTTTGTAGCTGACATGAGATGAGATCCGTGTGAGAGAGTCAATTT 8285  
QY 1325 ATCAATGTTGTAGCTTGGCCCCCGAAGCAGACAGGCTTAAGTGTCTACAGAGCGGC 1384  
DB 8286 ATCAATGTTGTAGCTTGGCCCCCGAAGCAGACAGGCTTAAGTGTCTACAGAGCGGC 8345



```

mat_peptide      7587. .9359      /product="NSSB"
misc_feature     1479. 1.559      /note="hypervariable region (HVR) "
misc_feature     6954. .7073      /note="interferon sensitivlty determining region (ISDR) "
N

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd

OM nucleic - nucleic search, using sw model

```
Run on: February 18, 2005, 20:55:25 ; Search time 990.878 Seconds
        (without alignments)
        10574.404 Million cell updates/sec
```

Title: US-09-664-363-3

Perfect score: 1770  
Sequence: 1 CAAATGACTTCCAGAGC.....ACTACTACCCGTGACCG 1770

Scoring table: IDENTITY\_NUC

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

### Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneeqn1980s: \*  
2: geneeqn1990s: \*  
3: geneeqn2000s: \*  
4: geneeqn2001as: \*  
5: geneeqn2001bs: \*  
6: geneeqn2002as: \*  
7: geneeqn2002bs: \*  
8: geneeqn2003as: \*  
9: geneeqn2003bs: \*  
10: geneeqn2003cs: \*  
11: geneeqn2003ds: \*  
12: geneeqn2004as: \*  
13: geneeqn2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1770	100.0	1770	2	AA012236	AA012236 Clone UG2
2	1762	99.5	3750	2	AA012241	AA012241 Encodes p
3	1553.8	87.8	8638	6	ABK88596	ABK88596 Hepatitis
4	1553.8	87.8	8638	6	ABK88578	ABK88578 Hepatitis
5	1553.4	87.8	8642	6	ABK88574	ABK88574 Hepatitis
6	1552.2	87.7	7141	6	AAD25333	AAD25333 Hepatitis
7	1552.2	87.7	7789	6	AAD25330	AAD25330 Hepatitis
8	1552.2	87.7	7987	6	AAD25329	AAD25329 Hepatitis
9	1552.2	87.7	7987	6	AAD25321	AAD25321 Hepatitis
10	1552.2	87.7	7987	6	AAD25324	AAD25324 Hepatitis
11	1552.2	87.7	7989	3	AAAS9968	AAAS9968 Hepatitis
12	1552.2	87.7	7989	6	AAD25322	AAD25322 Hepatitis
13	1552.2	87.7	7989	6	AAD25326	AAD25326 Hepatitis
14	1552.2	87.7	7989	6	AAD25325	AAD25325 Hepatitis
15	1552.2	87.7	7989	10	AAD93722	AAD93722 Hepatitis
16	1552.2	87.7	7989	10	AAD93724	AAD93724 Hepatitis
17	1552.2	87.7	7989	12	AD57885	AD57885 HCV repl
18	1552.2	87.7	7991	6	AA147279	AA147279 Hepatitis
19	1552.2	87.7	7992	6	AA147276	AA147276 Hepatitis
20	1552.2	87.7	7992	6	AA147280	AA147280 Hepatitis

21	1552.2	87.7	7992	6	AA1472777	AA1472777	Hepatitis
22	1552.2	87.7	7992	6	AA1472781	AA1472781	Hepatitis
23	1552.2	87.7	7995	6	AA1472778	AA1472778	Hepatitis
24	1552.2	87.7	8001	3	AA9898667	AA9898667	Hepatitis
25	1552.2	87.7	8637	3	AA9898967	AA9898966	Hepatitis
26	1552.2	87.7	8638	6	ABK868595	ABK868595	Hepatitis
27	1552.2	87.7	8638	6	ABK868577	ABK868577	Hepatitis
28	1552.2	87.7	8639	6	ABK868573	ABK868573	Hepatitis
29	1552.2	87.7	8643	6	ABK868575	ABK868575	Hepatitis
30	1552.2	87.7	8649	3	AA9898969	AA9898969	Hepatitis
31	1552.2	87.7	8732	10	ADE48121	ADE48121	HCV
32	1552.2	87.7	9605	6	ABK914331	ABK914331	Hepatitis
33	1552.2	87.7	9605	6	ABK914324	ABK914324	Hepatitis
34	1552.2	87.7	9605	6	ABK914329	ABK914329	Hepatitis
35	1552.2	87.7	9605	6	ABK914332	ABK914332	Hepatitis
36	1552.2	87.7	9605	6	ABK914311	ABK914311	Hepatitis
37	1552.2	87.7	9605	6	ABK914330	ABK914330	Hepatitis
38	1552.2	87.7	9605	6	ABK914328	ABK914328	Hepatitis
39	1552.2	87.7	9605	6	ABK914325	ABK914325	Hepatitis
40	1552.2	87.7	9605	6	ABK914326	ABK914326	Hepatitis
41	1552.2	87.7	9605	6	ABK914333	ABK914333	Hepatitis
42	1552.2	87.7	9605	6	AA025332	AA025332	Hepatitis
43	1552.2	87.7	9608	6	ABK914427	ABK914427	Hepatitis
44	1552.2	87.7	10690	6	ABK91448	ABK91448	Hepatitis
45	1552.2	87.7	10690	6	ABK91435	ABK91435	Hepatitis

## ALIGNMENTS

## RESULT 1

ID AAQ12236 standard; cDNA to rRNA; 1770 BP.

AC AAQ12236;

DT 25-MAR-2003 (revised)

XX

XX

XX

OS non-B hepati

PN GB2239245-A.

PD 26-JUN-1991

PF 17-DEC-1990; 90GB-00027250.

PR 18-DEC-1989; 89GB-00028562.

PR 03-MAR-1990; 90GB-00004814.

PA (WELL ) WELLCOME FOUND LTD

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

**XX**

DR P-PSDB; AAR12594.

Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA

2  
-  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
8

XX

CC isolated from serum of humans infected by the virus. Genomic RNA from

CC library in lambda gt11 which was screened with antisera and

CC oligonucleotide probes. One of the positive plaques to be obtained was  
 CC JG2. DNA was extracted from denatured phage stocks, amplified by PCR and  
 CC sequenced. See also AA012237-Q12242. (Updated on 25-MAR-2003 to correct  
 CC PA field.)

XX Sequence 1770 BP, 402 A, 531 C, 496 G, 341 T, 0 U, 0 Other;

Query Match 100.0%; Score 1770; DB 2; Length 1770;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAAAATGACTTCCCGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATAGATG 60
DB 1 CAAAATGACTTCCCGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGGATAGATG 60
QY 61 GCGGGGACATTTACCGCGGTGAGTCAGAGAACAGGTGATATCTGTGACCTTTCCAG 120
DB 61 GCGGGGACATTTACCGCGGTGAGTCAGAGAACAGGTGATATCTGTGACCTTTCCAG 120
QY 121 CCGCTCCGAGCGGAGGAGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTCTGGGAAA 180
DB 121 CCGCTCCGAGCGGAGGAGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTCTGGGAAA 180
QY 181 TCCAGAAATTTCCCAACAGCGATGCCCGCATGGGCAAGCCGGAATTACAACTCTCGCTG 240
DB 181 TCCAGAAATTTCCCAACAGCGATGCCCGCATGGGCAAGCCGGAATTACAACTCTCGCTG 240
QY 241 CTGAGATCTGGAGAGGCGCGGACCTACGTCCTCCAGGTGTATCATGGGTGCCCATGCGCA 300
DB 241 CTGAGATCTGGAGAGGCGCGGACCTACGTCCTCCAGGTGTATCATGGGTGCCCATGCGCA 300
QY 301 CCTACTAAGACCCCTCTATATACCACTCCACGAGAAAGAGACAGTTGTTCTGACAGAA 360
DB 301 CCTACTAAGACCCCTCTATATACCACTCCACGAGAAAGAGACAGTTGTTCTGACAGAA 360
QY 361 TTCACCGTGTCTTCTGCGCTTGGCGGAGCTTGGCCACAAAGGCTTTTGTGATGCTCCGACCG 420
DB 361 TTCACCGTGTCTTCTGCGCTTGGCGGAGCTTGGCCACAAAGGCTTTTGTGATGCTCCGACCG 420
QY 421 TCGGCGCTGACAGCGGCGACCGGCAACCGCCCTCCGACCAATCCCTCCGACGACGCGGCA 480
DB 421 TCGGCGCTGACAGCGGCGACCGGCAACCGCCCTCCGACCAATCCCTCCGACGACGCGGCA 480
QY 481 GCAAGATCTGACGTTGAGTGTATTCCTCCATGCCCCCTTTGAGGGGAGCCGCGGAGC 540
DB 481 GCAAGATCTGACGTTGAGTGTATTCCTCCATGCCCCCTTTGAGGGGAGCCGCGGAGC 540
QY 541 CCCGATCTCAGCGAGGGGTCTTGTCTACCGTGAAGTGAAGGCGGTCGACGATCGTC 600
DB 541 CCCGATCTCAGCGAGGGGTCTTGTCTACCGTGAAGTGAAGGCGGTCGACGATCGTC 600
QY 601 TGCTCTCGATGTCTTACCATGAGCAGGCGCTGATCAAGCCATGCGCTGGGAGAA 660
DB 601 TGCTCTCGATGTCTTACCATGAGCAGGCGCTGATCAAGCCATGCGCTGGGAGAA 660
QY 661 AGCAAGCTGCCCATCAAGCGGTTGAGCAACTTTGTGCGTCAACAACATGCTTAC 720
DB 661 AGCAAGCTGCCCATCAAGCGGTTGAGCAACTTTGTGCGTCAACAACATGCTTAC 720
QY 721 GCTACCAATCCCGAGCGGCAAGCCAGGCGGAGAAAGGTACCTTTGACAGACTGCA 780
DB 721 GCTACCAATCCCGAGCGGCAAGCCAGGCGGAGAAAGGTACCTTTGACAGACTGCA 780
QY 781 ATCTGAGCATCTACCAAGACGTCCTCAAGAGATGAAGCCAAAGCGTCCACAGTT 840
DB 781 ATCTGAGCATCTACCAAGACGTCCTCAAGAGATGAAGCCAAAGCGTCCACAGTT 840
QY 841 AAGGCTAAGCTTCTATCATGATAGAGAAAGCTGCAAGTCAAGCCCATTCCTGGCCAAA 900
DB 841 AAGGCTAAGCTTCTATCATGATAGAGAAAGCTGCAAGTCAAGCCCATTCCTGGCCAAA 900
QY 901 TCTAATTTGGGTATGGGCAAAAGAGCTCCGAACTATCCAGCAAGGCCATTACCAAC 960
DB 901 TCTAATTTGGGTATGGGCAAAAGAGCTCCGAACTATCCAGCAAGGCCATTACCAAC 960

```

```

DB 901 TCTAATTTGGGTATGGGCAAAAGAGCTCCGAACTATCCAGCAAGGCCATTACCAAC 960
QY 961 ATCCGCTCCGTGTGGGAGAGACTTTGTGGAAGACATTTGAAAACCAATTTGACCACTC 1020
DB 961 ATCCGCTCCGTGTGGGAGAGACTTTGTGGAAGACATTTGAAAACCAATTTGACCACTC 1020
QY 1021 ATGGCAAAAATGAGGTTTCTGCGTCCACAGAGAGAGAGCGCGCAAGCAGCTGCG 1080
DB 1021 ATGGCAAAAATGAGGTTTCTGCGTCCACAGAGAGAGAGCGCGCAAGCAGCTGCG 1080
QY 1081 CTATGTGTTCCTCCAGACTTGGGAGGTCGTGTGTGCGAATAATGCGCTCTATGACGTG 1140
DB 1081 CTATGTGTTCCTCCAGACTTGGGAGGTCGTGTGTGCGAATAATGCGCTCTATGACGTG 1140
QY 1141 GTCTCACCTCCCTCAAGCTGTGAATGGGCTCTCTGTACGATTCGATTAATTCCTGGA 1200
DB 1141 GTCTCACCTCCCTCAAGCTGTGAATGGGCTCTCTGTACGATTCGATTAATTCCTGGA 1200
QY 1201 CAGCGGATCGAGTTCCTGTGTAACGCTGGAATCAAGAAAGACCCCTATGGGCTTTGCA 1260
DB 1201 CAGCGGATCGAGTTCCTGTGTAACGCTGGAATCAAGAAAGACCCCTATGGGCTTTGCA 1260
QY 1261 TATGACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 1320
DB 1261 TATGACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 1320
QY 1321 ATTTATCAATGTGTGACTTGGGCGCGGACGAGCGCAATGAGTTCCTCAACAG 1380
DB 1321 ATTTATCAATGTGTGACTTGGGCGCGGACGAGCGCAATGAGTTCCTCAACAG 1380
QY 1381 CGGCTTTATATCGGAGGTCCTCCCTGACTAATTTCAAAAGGAGCAAGTGGGCTTATGCGCG 1440
DB 1381 CGGCTTTATATCGGAGGTCCTCCCTGACTAATTTCAAAAGGAGCAAGTGGGCTTATGCGCG 1440
QY 1441 TCGCGCGGAGCGGCTGTCTGACGACTGAGCTGCGGTAATCACTTCACTTTGTAAG 1500
DB 1441 TCGCGCGGAGCGGCTGTCTGACGACTGAGCTGCGGTAATCACTTCACTTTGTAAG 1500
QY 1501 GCGTCTGAGGCTGTGAGCTGCAAGGCTCCAGACTGCAAGTGTGTGGGAGAC 1560
DB 1501 GCGTCTGAGGCTGTGAGCTGCAAGGCTCCAGACTGCAAGTGTGTGGGAGAC 1560
QY 1561 GCGCTGTGATCTGTGAGAGCGCGGAAACCAAGAGAGAGCGGCGAGCTACGATG 1620
DB 1561 GCGCTGTGATCTGTGAGAGCGCGGAAACCAAGAGAGAGCGGCGAGCTACGATG 1620
QY 1621 TTCAAGAGGCTATGACTGATGACTCTGCCCCCGGGGACCCGCCAACAAGATAC 1680
DB 1621 TTCAAGAGGCTATGACTGATGACTCTGCCCCCGGGGACCCGCCAACAAGATAC 1680
QY 1681 GACCTGAGGTGATGAATCACTATGCTCCCAATGTGTGTGGGCGAGATGCTGGC 1740
DB 1681 GACCTGAGGTGATGAATCACTATGCTCCCAATGTGTGTGGGCGAGATGCTGGC 1740
QY 1741 AAAAGGATATCTACTCTCAACCCGTGACCCG 1770
DB 1741 AAAAGGATATCTACTCTCAACCCGTGACCCG 1770

```

RESULT 2  
 AA012241  
 ID AA012241 standard; cDNA, 3750 BP.  
 XX  
 XX AA012241;  
 XX AC  
 XX 25-MAR-2003 (revised)  
 DT 17-SEP-1991 (first entry)  
 XX  
 XX Encodes portion of PT-NANBH viral non-structural protein.  
 DE  
 XX post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.  
 KM  
 XX Non-A.  
 OS

OS non-B hepatitis virus.  
XX GB2239245-A.  
XX PD 26-JUN-1991.  
XX PF 17-DEC-1990; 90GB-00027250.  
XX PR 18-DEC-1989; 89GB-00028562.  
XX PR 27-FEB-1990; 90GB-0000414.  
XX PR 03-MAR-1990; 90GB-00004814.  
XX PA (WELL ) WELLCOME FOUND LTD.  
XX PA (HIGH/) HIGHFIELD P E.  
XX PI Highfield PE, Rodgers BC, Tedder RS, Barbara JAU;  
XX DR WPI; 1991-187584/26.  
XX DR P-PSDB; AARI2599.  
XX PT Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA  
XX PT and antibodies used in diagnostic assays and in vaccines.  
XX PS Claim 10; Page 88-97; 108pp; English.  
XX CC This sequence probably encodes viral non-structural proteins of the PT-  
CC NANBH viral genome which are antigenic. It was isolated from serum of  
CC humans infected by the virus. See also AAQ12236-40 and AAQ12242. (Updated  
CC on 25-MAR-2003 to correct PA field.)  
XX SQ Sequence 3750 BP; 794 A; 1140 C; 1072 G; 744 T; 0 U; 0 Other;  
Query Match 99.5%; Score 1762; DB 2; Length 3750;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1765; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 CAAATGACTTCCAGACGCTGACCTCATGAGGCACTCTGAGGCGGAGTGAAGATG 60  
DB 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCACTCTGAGGCGGAGTGAAGATG 2040  
QY 61 GGGGGGACATTAACCGCGTGAAGTCAAGAGCAAGAGTAACTCTGACTCTTTCGAC 120  
DB 2041 GGGGGGACATTAACCGCGTGAAGTCAAGAGCAAGAGTAACTCTGACTCTTTCGAC 2100  
QY 121 CGGCTCCGAGCGGAGAGTGAAGCGGAAAGTGTCCGTCGGGAGATCTCGCGAAA 180  
DB 2101 CGGCTCCGAGCGGAGAGTGAAGCGGAAAGTGTCCGTCGGGAGATCTCGCGAAA 2160  
QY 181 TCCAGAAATTCACAGGAGTCCCGCATGGGCAACGCCGATTAACAACCTCCGCTG 240  
DB 2161 TCCAGAAATTCACAGGAGTCCCGCATGGGCAACGCCGATTAACAACCTCCGCTG 2220  
QY 241 CTGGAATCTGGAAGCGCCGAGTACGTCCTTCAGTGTACATGGGTCCCACTGCCA 300  
DB 2221 CTGGAATCTGGAAGCGCCGAGTACGTCCTTCAGTGTACATGGGTCCCACTGCCA 2280  
QY 301 CCTACTAAGACCCCTCTCTATACCACTCCACGAGAAAGAGCAAGTTGTTGACGAA 360  
DB 2281 CCTACTAAGACCCCTCTCTATACCACTCCACGAGAAAGAGCAAGTTGTTGACGAA 2340  
QY 361 TCCACGATCTTCTGCGCTGGGAGCTTGGCAAAAGGCTTTGGTACTCCGACCG 420  
DB 2341 TCCACGATCTTCTGCGCTGGGAGCTTGGCAAAAGGCTTTGGTACTCCGACCG 2400  
QY 421 TCGGCGGTGACAGCGGACAGGCAACGCCCTCTGACCAATCTTCGACAGCGGCGA 480  
DB 2401 TCGGCGGTGACAGCGGACAGGCAACGCCCTCTGACCAACCTTCGACAGCGGCGA 2460  
QY 481 GCGAGATCTGACGTTGATGTTCTCTCATGCCCCCTTTGAGGGGAGCGGGGAGAC 540  
DB 2461 GCGAGATCTGACGTTGATGTTCTCTCATGCCCCCTTTGAGGGGAGCGGGGAGAC 2520  
QY 541 CCGAGATCTGACGAGCGGCTTGTGTCTACGATGATGAGGCGGCTGAGAGCTGTCTC 600

DB 2521 CCGAGATCTGACGAGCGGCTTGTGTCTACGATGATGAGAGCGGCTGAGAGCTGTCTC 2580  
QY 601 TGTGCTCATGATGTTCTTACATGACATGACAGCGGCTGTGATACCGCATGCGTGGAGGAA 660  
DB 2581 TGTGCTCATGATGTTCTTACATGACATGACAGCGGCTGTGATACCGCATGCGTGGAGGAA 2640  
QY 661 AGCAAGCTGCCATTCAGAGCGGTTGAGCAACTGTTGCTGCGTGCACCAACATGATCTAC 720  
DB 2641 AGCAAGCTGCCATTCAGAGCGGTTGAGCAACTGTTGCTGCGTGCACCAACATGATCTAC 2700  
QY 721 GCTACACATCCCGCAGCGCAAGCCAGCGGAGAGAAAGGTCACTTTGACAGACTGCA 780  
DB 2701 GCTACACATCCCGCAGCGCAAGCCAGCGGAGAGAAAGGTCACTTTGACAGACTGCA 2760  
QY 781 ATCTGAGAGATCACTTACAGAGAGCTGTCTCAAGAGATGAAGCGAGCGTCCACATTT 840  
DB 2761 ATCTGAGAGATCACTTACAGAGAGCTGTCTCAAGAGATGAAGCGAGCGTCCACATTT 2820  
QY 841 AAGGCTAAGCTTCTATCAGTAGAGAGCGCTGCAAGCTGACCGCCCACTTCGGCCAA 900  
DB 2821 AAGGCTAAGCTTCTATCAGTAGAGAGCGCTGCAAGCTGACCGCCCACTTCGGCCAA 2880  
QY 901 TCTAATTTGGCTATGAGGCAAGAGAGCTCCGAACTATCCAGCAAGGCACTTAACAC 960  
DB 2881 TCTAATTTGGCTATGAGGCAAGAGAGCTCCGAACTATCCAGCAAGGCACTTAACAC 2940  
QY 961 ATCCGCTCGGTGAGGAGACTTTGTTGGAAGACCTGAACCAATTGACACACCATC 1020  
DB 2941 ATCCGCTCGGTGAGGAGACTTTGTTGGAAGACCTGAACCAATTGACACACCATC 3000  
QY 1021 ATGGCAAAATGAGGTTTTTTCGCGTCAACAGAGAGAGAGCGGCAAGCGAGTCCG 1080  
DB 3001 ATGGCAAAATGAGGTTTTTTCGCGTCAACAGAGAGAGAGCGGCAAGCGAGTCCG 3060  
QY 1081 CTTATGCTTCCACAGACTTGGGAGGTCGCTGTGTGGAAGAAATGAGCCCTTATGACGTG 1140  
DB 3061 CTTATGCTTCCACAGACTTGGGAGGTCGCTGTGTGGAAGAAATGAGCCCTTATGACGTG 3120  
QY 1141 GTCTCAACCTCTCCAGGCTGTGATGAGCTCTCTCTGATTCAGATTCAGATTCCTGGA 1200  
DB 3121 GTCTCAACCTCTCCAGGCTGTGATGAGCTCTCTCTGATTCAGATTCAGATTCCTGGA 3180  
QY 1201 CAGCGGCTGAGGTTCTGTGTAAGCGCTGGAATCAAAAGCCCTATGAGCTTTGGA 1260  
DB 3181 CAGCGGCTGAGGTTCTGTGTAAGCGCTGGAATCAAAAGCCCTATGAGCTTTGGA 3240  
QY 1261 TATGACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 1320  
DB 3241 TATGACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCA 3300  
QY 1321 ATTATCAATGTTGTGACTTGGCCCGGAGCCAGACAGGAGCCATTAAGTCTGACAGAG 1380  
DB 3301 ATTATCAATGTTGTGACTTGGCCCGGAGCCAGACAGGAGCCATTAAGTCTGACAGAG 3360  
QY 1381 CGGCTTTATATCGGAGGTCCTCTGACTAATTCAAAAGGCAAGACTGCGGCTATCGCCGG 1440  
DB 3361 CGGCTTTATATCGGAGGTCCTCTGACTAATTCAAAAGGCAAGACTGCGGCTATCGCCGG 3420  
QY 1441 TGGCGGAGAGCGGCTGTGACAGTACTAGCTGAGTATACCTTCATGATCTTGAAG 1500  
DB 3421 TGGCGGAGAGCGGCTGTGACAGTACTAGCTGAGTATACCTTCATGATCTTGAAG 3480  
QY 1501 GCTCTGACAGCTGTGAGCTGCAAGAGCTCCAGAGCTGACAGATGCTGTGTGCGAGAC 1560  
DB 3481 GCTCTGACAGCTGTGAGCTGCAAGAGCTCCAGAGCTGACAGATGCTGTGTGCGAGAC 3540  
QY 1561 GGCCTTGTGTTATCTGTGAGAGCGGAGAAACCCAGAGAGAGCGCGGAGCCCTACAGATC 1620  
DB 3541 GGCCTTGTGTTATCTGTGAGAGCGGAGAAACCCAGAGAGAGCGCGGAGCCCTACAGATC 3600  
QY 1621 TTCAGGAGGCTATGACTAGTACTGTGCCCCCGGAGGAGCCGCGCAACAGAAATAC 1680



Db	3601	TTCAAGGAGCGCTATACATCAGGACCTCTGCCCCCCCCGGGGAGACCGGCCCAACCGAATATC	3660
Oy	1681	GACCTGAGTTGATTAACATCATGCTCTCCATCATGTGTGCGGCGACGATGATCTGGC	1740
Db	3661	GACCTGGAGTTGATTAACATCATGCTCTCCATCATGTGTGCGGCGACGATGATCTGGC	3720
Oy	1741	AAAAGGTAATCTACCTACCCCGTAGCCGG	1770
Db	3721	AAAAGGTAATCTACCTACCCCGTAGCCGG	3750
RESULT 3			
ID	ABK88596	ABK88596	
AC	ABK88596	standard; DNA; 8638 BP.	
XX			
DT	21-OCT-2002	(first entry)	
DE	Hepatitis C virus R3-rep-5'A replicon.		
XX			
KW	Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;		
KW	cell culture replication; ds; mutant.		
XX			
OS	Hepatitis C virus.		
XX			
Key	Location/Qualifiers		
FT	mutation	replace(1,G)	
FT		/tag= a	
FT	CDS	1802..8407	
FT		/tag= b	
FT		/product= "HCV NS2-5B"	
FT		/note= "Viral enzymes"	
FT	mutation	replace(2509,T)	
FT		/tag= c	
FT	mutation	replace(2778,G)	
FT		/tag= d	
FT	mutation	replace(2840,A)	
FT		/tag= e	
FT	mutation	replace(3574,T)	
FT		/tag= f	
FT	mutation	replace(4052,A)	
FT		/tag= g	
FT	mutation	replace(4387,C)	
FT		/tag= h	
FT	mutation	replace(4446,A)	
FT		/tag= i	
FT	mutation	replace(4507,C)	
FT		/tag= j	
FT	mutation	replace(4855,T)	
FT		/tag= k	
FT	mutation	replace(5351,A)	
FT		/tag= l	
FT	mutation	replace(5498,G)	
FT		/tag= m	
FT	mutation	replace(5659,G)	
FT		/tag= n	
FT	mutation	replace(5838,T)	
FT		/tag= o	
FT	mutation	replace(5871,C)	
FT		/tag= p	
FT	mutation	replace(6115,A)	
FT		/tag= q	
XX			
PN	WO200252015-A2.		
XX			
PD	04-JUL-2002.		
XX			
PF	20-DEC-2001; 2001WO-CA001843.		
XX			
PR	22-DEC-2000; 2000US-0257857P.		
XX			
PA	(BOEH ) BOEHRINGER INGELHEIM CANADA LTD.		

XX Kukolj G, Pause A;  
 P1  
 XX  
 XX  
 DR WPI; 2002-575382/61.  
 XX  
 XX  
 PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which  
 PT possess enhanced transduction or replication efficiency, useful for  
 PT evaluating potential inhibitors of HCV replication.  
 XX  
 XX Claim 11; Page 130-140; 140pp; English.

CC The invention describes a self-replicating hepatitis C virus (HCV)  
CC polynucleotide molecule comprising a 5'-non translated region (NTR),  
CC where guanine at position 1 is substituted for adenine, a HCV polyprotein  
CC region coding for a HCV polyprotein; and a 3'-NTR region. The self-  
CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating  
CC potential inhibitors of HCV replication. The HCV RNA molecule is also  
CC useful for efficiently establishing cell culture replication. The self-  
CC replicating polynucleotide molecule contains a 5'-NTR, where G at  
CC position 1 is substituted for A, and therefore provides an alternative to  
CC existing systems comprising a self-replicating HCV RNA molecule that, in  
CC conjunction with mutations in the HCV non-structural region, such as the  
CC G(2042)/C/R mutations, transduces and/or replicates with greater  
CC efficiency. This sequence represents hepatitis C virus replicon R3-rep-  
CC 5'G, a self-replicating HCV polynucleotide molecule created from the  
CC APGK12 (see ABK88573)

Sequence 8638 BP; 1767 A; 2558 C; 2440 G; 1873 T; 0 U; 0 Other;

Query Match 87.8%; Score 1553.8; DB 6; Length 8638;

Matches 1633; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

5 ATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCCTGTGGCGCATGAGATGGCG 64

Db 6030 ATGACTCCCGGACGCTGACCTCATCGAGGCCAACCTCCTGTGGCGGCAAGAGATGGCG 6081

65 GGGACATTACCCGGCTGGAGTCAGAGAACAGGTAGTAATCCTGGACTCTTTCGACCCGC 124

DB 6090 GGACATTCACCCCGCTGGAGTCAGAGAAAGGTAGTAAATTGGACTCTTTCGAGCCGC 6145

125 TCCAGCGAGGAGTGTCCGTCCCGCGAGATCTCGGAATCCA 184

DB 0130 LCAGGCGAGGAGH1GAGAGGAGH1M1CCG1LCGGGCGAGH1LC1GCGAGG1CM 0203

27

Q. Now, you're going to tell me that the defendant was not in the car at the time of the shooting, is that correct?

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 104

[illegible]

425 CCGTTCGACAGCGGCACCGCACCGCCCTCTGACCAATCCCTCCGACCGACCGCGAGCCAG 484

6450 CCGTCGACAGCGGCAACGGCCCTCTCTGACCAAGCCCTCCGACGACGGCGACGCGG 6500

485 GATCTGACCGTTGAGTGGTATTCCTCCATGCCCCCCCCCTTGAGGGGGAGCCGGGGGACCCCG 544

Db 6510 GATCCGACGTTGAGTCGTACTCTCCATGCCCCCTTGAGGGGGAGCCGGGGATCCCG 6565

545 ATCTCAGCGACGGTCTTGCTACCGTGAGTGAGGAGCCGGTGAGGACGTCGTCGCT 604

Db 6570 ATCTCAGCAGGGTCTTGGTCTACCGTAAGCAGGAGGCTAGTGAGGACGTCGTCTGCT 6629



QY 605 GCTCGATGTCCTACATGAGCAGCGGCTGTGATCAAGCCATGCGTGGGAGAAAGCA 664  
 DB 6630 GCTCGATGTCCTACATGAGCAGCGGCTGTGATCAAGCCATGCGTGGGAGAAAGCA 6689  
 QY 665 AGCTGCCATCAACGCGTTGAGCACTTTTGTGCTGCTACCAACATGGTCTAGCCTA 724  
 DB 6690 AGCTGCCATCAACGCGTTGAGCACTTTTGTGCTGCTACCAACATGGTCTAGCCTA 6749  
 QY 725 CCAATCCCGGAGCGGAGCGGAGCGGAGAAAGGTCACTTTGACAGACTGCAAAATCC 784  
 DB 6750 CCAATCCCGGAGCGGAGCGGAGCGGAGAAAGGTCACTTTGACAGACTGCAAAATCC 6809  
 QY 785 TGGAGCATACATCAAGGAGCTGTGCTCAAGGAGATGAAAGGAGGCTCCACAGTTAAAG 844  
 DB 6810 TGGAGCATACATCAAGGAGCTGTGCTCAAGGAGATGAAAGGAGGCTCCACAGTTAAAG 6869  
 QY 845 CTAACTTCTATCAGTATGAGAGAAAGCTGCAAGCTGACGCCCCACATTCGGCCAAATCTA 904  
 DB 6870 CTAACTTCTATCAGTATGAGAGAAAGCTGTAAGCTGACGCCCCACATTCGGCCAAATCTA 6929  
 QY 905 AATTGGCTATGGGGGCAAGAGAGCTCGGAACTTATCAGCAAGGCCATTAACCAATCC 964  
 DB 6930 AATTGGCTATGGGGGCAAGAGAGCTCGGAACTTATCAGCAAGGCCATTAACCAATCC 6989  
 QY 965 GCTCCGTTGGGAGGAGCTTTGTTGGAAGACATGAAACCAATTTGACACCAATCATG 1024  
 DB 6990 GCTCCGTTGGGAGGAGCTTTGTTGGAAGACATGAAACCAATTTGACACCAATCATG 7049  
 QY 1025 CAAAAATGAGATTTTCTGCGTCAACAGAGAGAGGCGGCAAGCCAGCTCGCCTTA 1084  
 DB 7050 CAAAAATGAGATTTTCTGCGTCAACAGAGAGAGGCGGCAAGCCAGCTCGCCTTA 7109  
 QY 1085 TCGTGTCCGAGACTTGGGGGCTCGTGTGTGCGAGAAATAGCCCTCTATGACGTGCT 1144  
 DB 7110 TCGTGTCCGAGACTTGGGGGCTCGTGTGTGCGAGAAATAGCCCTCTATGACGTGCT 7169  
 QY 1145 CCAACCTCCCTCAGGCTGTGATGAGGCTCCCTGATCCAGATTCAGATTCCTCGAGACG 1204  
 DB 7170 CCAACCTCCCTCAGGCTGTGATGAGGCTCCCTGATCCAGATTCAGATTCCTCGAGACG 7229  
 QY 1205 GGGTCAAGTCTGCTGTAACGCGCTGGAATCAAGAAAGACCCCTATGGGCTTTGCAATG 1264  
 DB 7230 GGGTCAAGTCTGCTGTAACGCGCTGTAACGCGCTGGAATCAAGAAAGACCCCTATGGGCTTTGCAATG 7289  
 QY 1265 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCGGTATGAGAGAGTCAATTT 1324  
 DB 7290 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCGGTATGAGAGAGTCAATTT 7249  
 QY 1325 ATGAATGTTGTGACTTGGCCCCCGAAGCGAGACAGGCGCATTAAGTGGCTACAGAGCGGC 1384  
 DB 7350 ATGAATGTTGTGACTTGGCCCCCGAAGCGAGACAGGCGCATTAAGTGGCTACAGAGCGGC 7409  
 QY 1385 TTTATATCGGGGGTCCCTGACTAATTCAAAGGGGAGAACTGCGGCTATCGCCGGTGC 1444  
 DB 7410 TTTATATCGGGGGTCCCTGACTAATTCAAAGGGGAGAACTGCGGCTATCGCCGGTGC 7469  
 QY 1445 GCGGAGCGGCTGTGAGCACTAGCTGCGGTATATACCTTCACATGTTACTTGAAGGCT 1504  
 DB 7470 GCGGAGCGGCTGTGAGCACTAGCTGCGGTATATACCTTCACATGTTACTTGAAGGCT 7529  
 QY 1505 CTGCAACCTGTGAGCTGCAAGCTCCAGAGCTGCAAGATGCTGTTGGGAGAGCGGC 1564  
 DB 7530 CTGCAACCTGTGAGCTGCAAGCTCCAGAGCTGCAAGATGCTGTTGGGAGAGCGGC 7589  
 QY 1565 TTGTCTTATCTGTGAGAGCGGAGAAACCCAGAGAGCGGCGAGCTTACGATCTTCA 1624  
 DB 7590 TTGTCTTATCTGTGAGAGCGGAGAAACCCAGAGAGCGGCGAGCTTACGATCTTCA 7649  
 QY 1625 CGGAGGCTATGACTAGTACTTGCCTCCCGGGAGCGGCGCCCAACGAAATACGAGC 1684  
 DB 7650 CGGAGGCTATGACTAGTACTTGCCTCCCGGGAGCGGCGCCCAACGAAATACGAGC 7709  
 QY 1685 TGGAGTTGATTAATCATCTGCTCCCAATGTGTGCGGAGAGATGATCTGGCAAAA 1744

DB 7710 TGGAGTTGATTAATCATCTGCTCCAAATGTGTGACTCGCGACATGATCTGGCAAAA 7769  
 QY 1745 GGGTATACTACTACCTACCCGCTGAGCC 1769  
 DB 7770 GGGTATACTACTACCTACCCGCTGAGCC 7794  
 RESULT 4  
 ABR88578  
 ID ABR88578 standard; DNA; 8638 BP.  
 XX  
 AC ABR88578;  
 XX  
 DT 21-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus R3-rep replicon.  
 XX  
 KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;  
 KW cell culture replication; ds; mutant.  
 XX  
 OS Hepatitis C virus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1802..8407  
 FT /tag= a  
 FT /product= "HCV NS2-5B"  
 FT /note= "Viral enzymes"  
 FT mutation replace(2509,T)  
 FT /tag= b  
 FT mutation replace(2778,G)  
 FT /tag= c  
 FT mutation replace(2840,A)  
 FT /tag= d  
 FT mutation replace(3574,T)  
 FT /tag= e  
 FT mutation replace(4052,A)  
 FT /tag= f  
 FT mutation replace(4387,C)  
 FT /tag= g  
 FT mutation replace(4446,A)  
 FT /tag= h  
 FT mutation replace(4507,C)  
 FT /tag= i  
 FT mutation replace(4855,T)  
 FT /tag= j  
 FT mutation replace(5351,A)  
 FT /tag= k  
 FT mutation replace(5498,G)  
 FT /tag= l  
 FT mutation replace(5659,G)  
 FT /tag= m  
 FT mutation replace(5838,T)  
 FT /tag= n  
 FT mutation replace(5871,C)  
 FT /tag= o  
 FT mutation replace(6115,A)  
 FT /tag= p  
 PN WO200252015-A2.  
 PD 04-JUL-2002.  
 PF 20-DEC-2001; 2001WO-CA001843.  
 PR 22-DEC-2000; 2000US-0257857P.  
 PA (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.  
 PI Kukulj G, Paube A;  
 XX WPI; 2002-575382/51.



ID	Accession	Source	Length	Score	DB	Length
DB	7770	GGGTGACTATCTCACCCGTGACC	7794			
RESULT 5						
ABK88574						
ID	ABK88574	standard; DNA; 8642 BP.				
XX						
AC	ABK88574;					
XX						
DT	21-OCT-2002 (first entry)					
XX						
DE	Hepatitis C virus S 22-3 replicon.					
XX						
KM	Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;					
XX	cell culture replication; ds; mutant.					
XX						
OS	Hepatitis C virus.					
XX						
PH	Key	Location/Qualifiers				
FT	mutation	replace(1,G)				
FT		/*tag= a				
FT		1802. .8407				
FT	CDS	/*tag= b				
FT		/product= "HCV NS2-5B"				
FT		/note= "viral enzymes"				
FT	mutation	replace(4446,A)				
FT		/*tag= c				
FT	mutation	replace(5498,G)				
FT		/*tag= d				
FT	mutation	replace(6268,A)				
FT		/*tag= e				
XX						
FN	WO200252015+A2.					
XX						
PD	04-JUL-2002.					
XX						
PE	20-DEC-2001; 2001WO-CA001843.					
XX						
PR	22-DEC-2000; 2000US-0257857P.					
XX						
PA	(BOEH ) BOEHRINGER INGELHEIM CANADA LTD.					
PI	Kukolj G, Pause A;					
DR	WPI; 2002-575382/61.					
DR	P-PSDB; ABG30582.					
XX						
PT	New self-replicating RNA molecules from Hepatitis C virus (HCV), which					
PT	possess enhanced transduction or replication efficiency, useful for					
PT	evaluating potential inhibitors of HCV replication.					
PS	Claim 11; Page 59-69; 140pp; English.					
XX						
CC	The invention describes a self-replicating hepatitis C virus (HCV)					
CC	polynucleotide molecule comprising a 5'-non translated region (NTR),					
CC	where guanine at position 1 is substituted for adenine, a HCV polyprotein					
CC	region coding for a HCV polyprotein; and a 3'-NTR region. The self-					
CC	replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating					
CC	potential inhibitors of HCV replication. The HCV RNA molecule is also					
CC	useful for efficiently establishing cell culture replication. The self-					
CC	replicating polynucleotide molecule contains a 5'-NTR, where G at					
CC	position 1 is substituted for A, and therefore provides an alternative to					
CC	existing systems comprising a self-replicating HCV RNA molecule that, in					
CC	conjunction with mutations in the HCV non-structural region, such as the					
CC	G(2042)C/R mutations, transduces and/or replicates with greater					
CC	efficiency. This sequence represents hepatitis C virus replicon S22-3, a					
CC	self-replicating HCV polynucleotide molecule created from the replicon					
CC	APK12 (see ABK88753)					
XX						
Sequence	8642 BP; 1768 A; 2557 C; 2438 G; 1877 T; 0 U; 2 Other;					
Query Match	87.8%; Score 1553.4; DB 6; Length 8642;					
Best Local Similarity	.92.5%; Ptd No. 0;					

Matches 1632: Conservative				1;	Mismatches 132;	Indels	0;	Gaps	0;
QY	5	ATGAC	TTCCAGACGCTGACCTCATTCGAGGCCAACTCTGTGGCGCATAGATGGCG	64					
Db	6030	ATGACT	CCCCCGAGCGCTGACCTCATTCGAGGCGCAACTCTGTGGCGCAGAGATGGCGCG	6085					
QY	65	GGGAC	ATTACCCCGGTGGAGTCAAGAAACAAGTAACTCTGGACTCTTTCCAGCCCG	124					
Db	6090	GGAA	CATCACCCCGGTGGAGTCAAAAAATGAAGTAAATTTTGGACTCTTTCCAGCGCG	6145					
QY	125	TCCGAG	CGAGAGATGAGCGGGAGGTCCGTCGCGCGAGATCTCTGCGGAAATCCA	184					
Db	6150	TCCAAG	CGAGAGAGATGAGGAGATATCGAGGAGATTCCTTCGCGGAGATCTCTGCGAGGTCCA	6205					
QY	185	AGAAAT	TTCCACCAAGCATGCGCATGAGCACGCCGGATTACACCCTCGCTCTGG	244					
Db	6210	GGAAAT	TCTCCGAGCATGCGCATATGAGGACGCCGGATTACACCCTCCACTGTTTGG	6265					
QY	245	AGTCT	TGGAAAGCCCCCGGACTACGTCTCTCCAGTGGTACATGGGTGCCCACTGCCACTA	304					
Db	6270	AGTCT	TGGAAAGACCCCGGACTACGTCTCTCCAGTGGTACACGGGTGTCCATTGCGCCTG	6322					
QY	305	CTAAG	ACCCCTCTATACACCTCCACGAGAAAGAGGACAGTTGTTCTGACAGATCCA	364					
Db	6330	CCAAAG	CCCCCTCCCATACCACTCCACGAGAAAGAGGACGTTGTCTCTGACAGATCTA	6385					
QY	365	CCGTGT	CTTTCGCCCTGGCGGAGCTTGGCAAAAGCTTTTGTATGCTCCGAGCCGTGG	424					
Db	6390	CCGTGT	CTTTCGCCCTGGCGGAGCTTGGCAAAAGACTTTCGGGACGTCGGAATCTGTGG	6445					
QY	425	CCGTG	CACAGCGGCAACCGCCCTCTGACCAATCTCCGACGACGCGCGAGGAC	484					
Db	6450	CCGTG	CACAGCGGCAACCGCCCTCTCTGACCAACGCTCCGACGACGCGCAACGCGG	6505					
QY	485	GATCG	ACGTGATTCGATTCGATTCGTCAGTACCCCTTGAGAGGAGAGCGCGGAGCCCG	544					
Db	6510	GATCG	ACGATTTGATTCGATTCGTCAGTACCCCTTTGAGAGGAGAGCGCGGAGATCCCG	6565					
QY	545	ATCTC	AGCGAGCGGCTCTTGTCTTACCGTGAATGAGAGCGCGATGAGGACGTCTGCT	604					
Db	6570	ATCTC	AGCGAGCGGCTCTTGTGTCTACCGTGAAGCGAGAGGCTTATGAGGACGTCTGCT	6622					
QY	605	GCTCG	ATGTCTTACATGACAGGCGCTCTGATCAGCCATGCGCTGCGAGGAAACA	664					
Db	6630	GCTCG	ATGTCTCTTACATGACAGGCGCCCTGATCAGCCATGCGCTGCGAGGAAACA	6685					
QY	665	AGCGG	CCCATCAAGCGCTTGAACACTTTTGTGCGGTCAACAACAATGGTCTACGTGA	724					
Db	6690	AGCGG	CCCATCAATGACACTGAGCAACTTTTGTCTCCGTCAACAACACTTGTCTATGTGA	6745					
QY	725	CCACAT	CCCGAGCGCAAGCGAGCGGAGAGAGATCACTTTGACAGATCGCAAAATCC	784					
Db	6750	CAACAT	CTCGAGCGCAAGCCTGCGGAGAGAGATCACTTTGACAGATCTGCAAGTCC	6805					
QY	785	TGACG	ATCACTACAGACGTGCTCAAGAGATGAAGCGAGGCGTCCACAGTTAAAG	844					
Db	6810	TGACG	ACCACTACCGGGACGTGCTCAAGAGATGAAGCGAGGCGTCCACAGTTAAAG	6865					
QY	845	CTAAG	CTTATCATGTAAGAAACCTTGAAGCCGCCCAATTTGCGCCAAATCTA	904					
Db	6870	CTTAA	CTTCTATCCGTGAGAGAACTCTGTAAAGCTTGAAGCCCCCAATTTGCGCCAGATCTA	6925					
QY	905	AATTTG	CTATGAGGCAAGACGTCCGGAACCTATCCAGCAAGGCAATTAAACAATCC	964					
Db	6930	AATTTG	CTATGAGGCAAGACCTCCGGAACCTATCCAGCAAGGCGCTTAAACAATCC	6985					
QY	965	GCTCC	GTTGGAGACTTGTGGAGACACTGAACAACAATTTGACACCAATCATGG	1024					
Db	6990	GCTCC	GTTGGAGACTTGTCTGGAGACACTGAGACCAACAATTTGACACCAATCATGG	7045					
QY	1025	CAAAAA	TGAGGTTTTTCTGCGTCCACAGAGAGAGGCGCGGAACCGAGTGGCCTTA	1084					
Db	7050	CAAAAA	TGAGGTTTTTCTGCGTCCACAGAGAGAGGCGCGGAACCGAGTGGCCTTA	7105					

```
OY 1085 TCGTGTTCACAGACTTGGGGGTCCTGTGTGCGAGAAATGGCCCTTAATGACGTGTCT 1144
XX |||
XX |||
XX |||
Db 7110 TCGTATTCACAGATTGGGGGTTCTGTGTGCGAGAAATGGCCCTTAATGATGTGTCT 7169
OY 1145 CCACCTCCCTCAGGCTGTGATGGGCTCTGTAGGATTCAGATATTCCTCTGACACG 1204
XX |||
XX |||
XX |||
Db 7170 CCACCTCCCTCAGGCGGTGATGGGCTCTGTATAGGATTCAGATATTCCTCTGACACG 7229
OY 1205 GGGTGAAGTCTGTGGAACGCGCTGGAATCAAGAGAGACCCCTATGGGCTTTGATATG 1264
XX |||
XX |||
XX |||
Db 7230 GGGTGAAGTCTGTGGAATGCTGGAAGCAAGAAATGCCCTATGGGCTTTGATATG 7289
OY 1265 ACACCCGCTGTTTGAATCAACAGTCACTAGAGATGACATCCGCTGAGAGATCAATTT 1324
XX |||
XX |||
XX |||
Db 7280 ACACCCGCTGTTTGAATCAACAGTCACTAGAGATGACATCCGCTGAGAGATCAATTT 7349
OY 1325 ATCAATGTTGTACTTGGGCCCCGGAAGCCAGACAGGCTTAAGTGTGCTCAGACAGCGC 1384
XX |||
XX |||
XX |||
Db 7350 ACCAATGTTGTACTTGGGCCCCGGAAGCCAGACAGGCTTAAGTGTGCTCAGACAGCGC 7409
OY 1385 TTATATCGGGGGGTCCTGACTAATTTCAAAAGGAGAACTGCGGGCTATGCGGGTGC 1444
XX |||
XX |||
XX |||
Db 7410 TTATATCGGGGGGTCCTGACTAATTTCAAAAGGAGAACTGCGGGCTATGCGGGTGC 7469
OY 1445 GCGGAGCGGCGTGTGACGACTAGCTGCGGTATATACCTCAATGTTACTTGAAGGCT 1504
XX |||
XX |||
XX |||
Db 7470 GCGGAGCGGCGTGTGACGACTAGCTGCGGTATATACCTCAATGTTACTTGAAGGCT 7529
OY 1505 CTGACGCTGTGTGACGCTGCAAAAGCTTCAGAGATTCACGATGCTCGTGTGCGAGACGCGC 1564
XX |||
XX |||
XX |||
Db 7530 CTGACGCTGTGTGACGCTGCAAAAGCTTCAGAGATTCACGATGCTCGTGTGCGAGACGCGC 7589
OY 1565 TTGTGCTTATCTGTGAGAGCGGGGAAACCAGAGAGAGCGCGGAGCTCAGAGTCTCA 1624
XX |||
XX |||
XX |||
Db 7590 TTGTGCTTATCTGTGAGAGCGGGGAAACCAGAGAGAGCGCGGAGCTCAGAGTCTCA 7649
OY 1625 CGGAGGCTATGACTAGTACTGTGCCCCCGGGGAGACCCGCCACAGCAATATGACAC 1684
XX |||
XX |||
XX |||
Db 7650 CGGAGGCTATGACTAGTACTGTGCCCCCGGGGAGACCCGCCACAGCAATATGACAC 7709
OY 1685 TGGAGTGTATACATCATGCTCTCTCCATGTGTCCGTGCGGACGATGATCTGGCAAAA 1744
XX |||
XX |||
XX |||
Db 7710 TGGAGTGTATACATCATGCTCTCTCCATGTGTGTGATGCTGCGGACGATGATCTGGCAAAA 7769
OY 1745 GGGTATTAATCACTCAAGCCGCTGACCC 1769
XX |||
XX |||
XX |||
Db 7770 GGGTATTAATCACTCAAGCCGCTGACCC 7794
XX |||
XX |||
XX |||
RESULT 6
AAD25333
ID AAD25333 standard; cDNA; 7141 BP.
XX
AC AAD25333;
XX
DT 12-MAR-2002 (first entry)
XX
DE Hepatitis C virus (HCV) adaptive replicon 5'NTR-EMCV/HCVrepVII cDNA.
XX
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KW EMCV/HCVrepVII; ss.
XX
OS Hepatitis C virus.
XX
PN WO200189364-A2.
XX
29-NOV-2001.
PD
XX
23-MAY-2001; 2001MO-US016822.
PF
XX
23-MAY-2000; 2000US-00576989.
PR
```

```
XX (UNIM ) UNIV WASHINGTON.
PA
XX
XX Rice CM, Blight KJ;
PI
XX
XX WPI: 2002-066755/09.
DR
XX
XX Hepatitis C virus variants having greater transfection efficiency and
XX ability to survive subpassage, useful as a vaccine for immunizing primate
XX to the virus, comprise non-naturally occurring viral sequences.
XX
XX Example 2; Page 100-102; 174dp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
XX polynucleotides comprising non-naturally occurring HCV sequence and HCV
XX variants that have a transfection efficiency and ability to survive
XX subpassage greater than HCV that have wild-type polypeptide coding
XX regions. The polynucleotides of the invention are useful for identifying
XX a cell line that is permissive for infection with HCV and detecting
XX replication of HCV in cells of the cell line. They are also useful for
XX testing a compound for anti-viral properties and for inhibiting HCV
XX infection. They are also useful for the generation of defined HCV virus
XX stocks to develop in vitro and in vivo assays for virus neutralisation,
XX attachment, penetration and entry, structure/function studies on HCV
XX proteins and RNA elements and identification of new antiviral targets, a
XX systematic survey of cell culture systems and conditions to identify
XX those that support wild-type and variant HCV RNA replication and particle
XX release, production of adaptive HCV variants capable of more efficiency
XX replication in cell culture, production of HCV variants with altered
XX tissue or species tropism, establishment of alternative animal models for
XX inhibitor evaluation including those supporting HCV variant replication,
XX development of cell-free HCV replication assays, production of
XX immunogenic HCV particles for vaccination, engineering of attenuated HCV
XX derivatives as possible vaccine candidates, engineering of attenuated HCV
XX defective HCV derivatives for expression of heterologous gene products
XX for gene therapy and vaccine applications and for utilisation of the HCV
XX glycoproteins for targeted delivery of therapeutic agents to the liver
XX or other cell types with appropriate receptors. Vaccine comprising these
XX sequences is useful for inducing immunoprotection to HCV in a primate.
XX The present sequence is Hepatitis C virus (HCV) adaptive replicon 5'NTR-
XX EMCV/HCVrepVII cDNA
XX
SQ Sequence 7141 BP; 1481 A; 2130 C; 1982 G; 1548 T; 0 U; 0 Other;
XX
XX Query Match 87.7%; Score 1552.2; DB 6; Length 7141;
XX Best Local Similarity 92.5%; Pred. No. 0;
XX Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
OY 5 ATGACTTCCAGAGCGCTGACCTCATGAGGCCAACTTCCTGTGCGGCGATGATGGCG 64
XX |||
XX |||
XX |||
Db 4533 ATGACTCCCGGAGCGTGAAGCTCATGAGGCCAACTTCCTGTGCGGCGATGATGGCG 4592
OY 65 GGGACATTACCCGGGTGTGAGTCAAGAAAGATATCTCTGATCTTTGACACCCG 124
XX |||
XX |||
XX |||
Db 4593 GGAACATCACCCGGGTGTGAGTCAAGAAAGATATCTCTGATCTTTGACACCCG 4652
OY 125 TCCGAGCGAGAGAGATGAGCGGAGTGTCCGTCGCGGAGATCTCTGCGAAATCCA 184
XX |||
XX |||
XX |||
Db 4653 TCCAGCGGAGAGAGATGAGCGGAGTGTCCGTCGCGGAGATCTCTGCGAGATCCA 4712
OY 185 AGAATATCCACAGAGATGCGCGATGGGACGCCCGGATTAACACCTCCGCTGTG 244
XX |||
XX |||
XX |||
Db 4713 GGAATATCCCGTGAAGATGCGCGATGGGACGCCCGGATTAACACCTCCGCTGTG 4772
OY 245 AGTCTGGAAGGCCCCGAGCTAGTCTCTCCAGTGTGATCATGGGTGCCCACTGCACTTA 304
XX |||
XX |||
XX |||
Db 4773 AGTCTGGAAGGCCCCGAGCTAGTCTCTCCAGTGTGATCATGGGTGCCCACTGCACTTA 4832
OY 305 CTAAGACCCCTCTATACCACTTCACGAGAGAAAGAGACAGTGTGTCAGACAGATCCA 364
XX |||
XX |||
XX |||
Db 4833 CTAAGACCCCTCTCTCATACCTTCACGAGAGAAAGAGACAGTGTGTCAGACAGATCCA 4892
OY 365 CCGTGTCTTGTGCCCTGCGGAGCTTGCACAAAGGCTTTGTGTAGCTCCGGAACCTGCG 424
XX |||
XX |||
XX |||
```

```

Db      4893  CCGTGTCTTCTGCTGGGGAGCTGCCAAGAAAGCTTCGGAGCTCCGAAATGCTCGG 4952
Qy      425  CCGTCAGACGGGCGACGAGCAACGGCCCTCTCTGACCAATCTTCGACAGACGGGAGCAG 484
Db      4953  CCGTCAGACGGGCGACGAGCAACGGCCCTCTCTGACCAAGCTTCGACAGACGGGAGCAG 5012
Qy      485  GATCTGACGCTGAGTGTGTAATTCCTCATGCCCCCTTGAAGGGGGAGCCGGGGAGCCCG 544
Db      5013  GATTCGACGCTGAGTGTGTAATTCCTCATGCCCCCTTGAAGGGGGAGCCGGGGAGTCCG 5072
Qy      545  ATCTCAGCAGCGGCTCTTGTGCTACCGTGAAGAGAGCGGCTGAAGAGCTGTCTGCT 604
Db      5073  ATCTCAGCAGCGGCTCTTGTGCTACCGTGAAGAGAGCGGCTGAAGAGCTGTCTGCT 5132
Qy      605  GCTCGATGCTCTACATAGAGAGCGGCTCTGATCAAGCCATGCGCTGGAGAGAAAGCA 664
Db      5133  GCTCGATGCTCTACATAGAGAGCGGCTCTGATCAAGCCATGCGCTGGAGAGAAAGCA 5192
Qy      665  AGCTGCCATCAACGGGCTTGAAGCACTTTGCTGGTCAACCATGCTCTAGGCTA 724
Db      5193  AGCTGCCATCAACGGGCTTGAAGCACTTTGCTGGTCAACCATGCTCTAGGCTA 5252
Qy      725  CCACATCCCGCAGCGCAGACGCGCAGAGAGAGGTCACCTTTGACAGACTGCAATCC 784
Db      5253  CAACATCTCGCAGCGCAGACGCGCAGAGAGAGGTCACCTTTGACAGACTGCAAGTCC 5312
Qy      785  TGGAGCATCACTACAGAGACGTGCTCAAGAGATGAAGGGGAGGCGCTCAAGTTAAG 844
Db      5313  TGGAGCATCACTACAGAGACGTGCTCAAGAGATGAAGGGGAGGCGCTCAAGTTAAG 5372
Qy      845  CTAAAGCTTATACAGAGAGAGAGGCTGAGAGCTGACGCCCCACATTTGGGCCAATCTA 904
Db      5373  CTAAAGCTTATACAGAGAGAGAGGCTGAGAGCTGAGAGCTGAGGCCCCACATTTGGGCCAATCTA 5432
Qy      905  AATTGGCTATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 964
Db      5433  AATTGGCTATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5492
Qy      965  GCTCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1024
Db      5493  GCTCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5552
Qy      1025  CAAGAAATGAGGTTTCTGCTGCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084
Db      5553  CAAGAAATGAGGTTTCTGCTGCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5612
Qy      1085  TCGTGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1144
Db      5613  TCGTGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5672
Qy      1145  CCACCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1204
Db      5673  CCACCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5732
Qy      1205  GGGTGAAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1264
Db      5733  GGGTGAAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5792
Qy      1265  ACACCCGCTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1324
Db      5793  ACACCCGCTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5852
Qy      1325  ATCAATGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1384
Db      5853  ACCAATGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5912
Qy      1385  TTTATATGAGGAGTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1444
Db      5913  TTTATATGAGGAGTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5972
Qy      1445  GGGCAGAGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1504

```

```

Db      5973  GCGCAGCGGCTGATGACAGACAGCTGGGTAATACCTTCATATGTTGAAGGCGG 6032
Qy      1505  CTGACAGCTGTGAGAGTGAAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1564
Db      6033  CTGACAGCTGTGAGAGTGAAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6092
Qy      1565  TTGCTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1624
Db      6093  TTGCTGTTATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6152
Qy      1625  CGAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1684
Db      6153  CGAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6212
Qy      1685  TGGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1744
Db      6213  TGGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6272
Qy      1745  GGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1769
Db      6273  GGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6297

RESULT 7
AAD25330
ID  AAD25330 standard; cDNA; 7789 BP.
XX
AC  AAD25330;
XX
DT  12-MAR-2002 (first entry)
XX
DE  Hepatitis C virus (HCV) adaptive replicon HCVrep/NS2-5B cDNA.
XX
KW  Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW  gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KW  adaptive replicon; HCVrep/NS2-5B; ss.
XX
OS  Hepatitis C virus.
XX
PN  WO200189364-A2.
XX
PD  29-NOV-2001.
XX
PF  23-MAY-2001; 2001WO-US016822.
XX
PR  23-MAY-2000; 2000US-00576989.
XX
PA  (UNIW ) UNIV WASHINGTON.
XX
PI  Rice CM, Blight KJ;
XX
DR  WPI; 2002-066755/09.
XX
PT  Hepatitis C virus variants having greater transfection efficiency and
PT  ability to survive subpassage, useful as a vaccine for immunizing primate
PT  to the virus, comprise non-naturally occurring viral sequences.
XX
PS  Example 2; Page 90-92; 174pp; English.
XX
CC  The invention relates to Hepatitis C virus (HCV) variants which include
CC  polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC  variants that have a transfection efficiency and ability to survive
CC  subpassage greater than HCV that have wild-type polypeptide coding
CC  regions. The polynucleotides of the invention are useful for identifying
CC  a cell line that is permissive for infection with HCV and detecting
CC  replication of HCV in cells of the cell line. They are also useful for
CC  testing a compound for anti-viral properties and for inhibiting HCV
CC  infection. They are also useful for the generation of defined HCV virus
CC  stocks to develop in vitro and in vivo assays for virus neutralisation,
CC  attachment, penetration and entry, structure/function studies on HCV
CC  proteins and RNA elements and identification of new antiviral targets, a
CC  systematic survey of cell culture systems and conditions to identify
CC  those that support wild-type and variant HCV RNA replication and particle

```

CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of attenuated or  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilisation of the HCV  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
CC The present sequence is Hepatitis C virus (HCV) adaptive replicon  
CC HCVrep/NS2-5B CDNA  
XX  
SQ

Sequence 7789 BP; 1602 A; 2320 C; 2175 G; 1692 T; 0 U; 0 Other;

Query Match 87.7%; Score 1552.2; DB 6; Length 7789;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```
QY 5 ATGACTTCCCAAGCGTGAACCTCATCGAGGCCAACCTCTGTGGGGCATGAGTGGCG 64
DB 5181 ATGACTTCCCGAGACGCTGAACCTCATCGAGGCCAACCTCTGTGGGGCATGAGTGGCG 5240
QY 65 GGGACATTACCCCGGTGAGTCAAGAACAGATAGTATCTTGAATCTTTTCAGACCGC 124
DB 5241 GGAACATCACCCGCGTGAATGAGAAATAGATAGTATTTTGAATCTTTTCAGACCGC 5300
QY 125 TCCGAGCGGAGAGATGACGGGAAAGTCCGTCCCGCGGAGATCTTCGGAATATCCA 184
DB 5301 TCCAGCGGAGAGATGAGAGGAAAGTATCCGTTCCGCGGAGATCTTCGGAATATCCA 5360
QY 185 AGAATTTCCACAGAGATGCCGATGGGCAAGCCGGATTTACAACCTCCGCTGGC 244
DB 5361 GGAATTTCCCTGAGCGATGCCATATGGGCAAGCCGGATTTACAACCTCCGCTGGC 5420
QY 245 AGTCTGGAAGGCGCCGGAATGATGCTCCCTCAAGTGTATGATGGTCCCACTGCCACTA 304
DB 5421 AGTCTGGAAGGACCCGGAATGATGCTCCCTCAAGTGTATGATGGTCCCACTGCCACTA 5480
QY 305 CTAAACACCCCTCTATACCACTTCAACGGAAGAAAGACAGTGTGTTGACAGATCCA 364
DB 5481 CCAAGCCCTCTCGATACCACTTCAACGGAAGAAAGACAGTGTGTTGACAGATCCA 5540
QY 365 CCGTGTCTTCTGCGCGGAGGCTTGCACAAAGGCTTTGGTATGCTCCGGAACGCTGG 424
DB 5541 CCGTGTCTTCTGCGCGGAGGCTTGCACAAAGGCTTTGGTATGCTCCGGAACGCTGG 5600
QY 425 CCGTGAACGCGGACAGGCAACGCGCCCTCTGACCAATCTTCGACGACGCGGAGAG 484
DB 5601 CCGTGAACGCGGACAGGCAACGCGCTCTCTGACCAAGCTTCGACGACGCGGAGAG 5660
QY 485 GATTCGACGTTGATGCTATTTCTCCATGACCCCTTGAAGGGAGAGCGCGGGAGCCCG 544
DB 5661 GATTCGACGTTGATGCTATTTCTCCATGACCCCTTGAAGGGAGAGCGCGGGAGCTCCG 5720
QY 545 ATTCAGGAGGAGGCTTGTGTACGCTAGAGGAGGAGCGGAGAGAGAGAGAGAGAGAG 604
DB 5721 ATTCAGGAGGAGGCTTGTGTGTACGCTAGAGGAGGAGAGAGAGAGAGAGAGAGAG 5780
QY 605 GCTCGATGCTTACATGAGCAGGCGCTGTGATCAGCGCATGCGCTCGAGAGAGAGCA 664
DB 5781 GCTCGATGCTTACATGAGCAGGCGCTGTGATCAGCGCATGCGCTCGAGAGAGAGCA 5840
QY 665 AGCTGCCATCAACGCGTTGAGCAACTTTTCTGCTGCTACCAACAATGCTTACGCTA 724
DB 5841 AGCTGCCATCAATCACTGAGCAACTTTTCTGCTGCTACCAACAATGCTTACGCTA 5900
QY 725 CCACATCCCGAGCGCAAGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
DB 5901 CCACATCTCGAGCGCAAGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5960
```

```
QY 785 TGGAGATCACTACGAGACGTGTCTCAAGAGATGAAGAGCGGCTCCACAGTTAAG 844
DB 5961 TGGAGATCACTACGAGACGTGTCTCAAGAGATGAAGAGCGGCTCCACAGTTAAG 6020
QY 845 CTAACTTCTATCACTAAGAGAGAGCTTGAAGCTGACCGCCACATTTGGCGCAATCTA 904
DB 6021 CTAACTTCTATCACTAAGAGAGAGCTTGAAGCTGACCGCCACATTTGGCGCAATCTA 6080
QY 905 AATTGGCTAATGGGCAAGAGACGTCGGAAACCTATCCAGAGAGGCGATTAAACCATCC 964
DB 6081 AATTGGCTAATGGGCAAGAGACGTCGGAAACCTATCCAGAGAGGCGATTAAACCATCC 6140
QY 965 GCTCCGTGAGAGAGACTTGTGGAAGACACTGAAACCAATTTGACACCAACCATCATG 1024
DB 6141 GCTCCGTGAGAGAGACTTGTGGAAGACACTGAAAGACCAATTTGACACCAACCATCATG 6200
QY 1025 CAAAAATGAGGTTTCTGCTCCAAACGAGAGAGAGAGGCGGCAAGCAGCTGCTTTA 1084
DB 6201 CAAAAATGAGGTTTCTGCTCCAAACGAGAGAGAGGCGGCAAGCAGCTGCTTTA 6260
QY 1085 TCGTGTCCAGACTTGGGGGTCGCTGTGTGTCGAGAAATGAGCCCTATGACGTGTCT 1144
DB 6261 TCGTGTCCAGACTTGGGGGTCGCTGTGTGTCGAGAAATGAGCCCTATGACGTGTCT 6320
QY 1145 CCACCTTCCCTCAGGCTGTGATGGGCTCTCTGTAACGATTCAGTATTTCTCTGACAGC 1204
DB 6321 CCACCTTCCCTCAGGCGGTGATGGGCTCTCTGTAACGATTCAGTATTTCTCTGACAGC 6380
QY 1205 GGGTGAATTTCTGTGTAACGCTTGAAGATTAAGAAAGACCCCTATGGGCTTTGCTATG 1264
DB 6381 GGGTGAATTTCTGTGTATCTCTGTAAGCTTGAAGAAATGCTCTATGGGCTTTGCTATG 6440
QY 1265 AACCCGCTTTTGTGACTCAACAGTCACTGAGATGATCCGCTGTGAGAGAGTCAATTT 1324
DB 6441 AACCCGCTTTTGTGACTCAACAGTCACTGAGATGATCCGCTGTGAGAGAGTCAATTT 6500
QY 1325 ATCAATGTTGTGACTTGGCGCCCGGAGAGCAGAGCCATTAAGTCTCTCAAGAGCGGC 1384
DB 6501 ATCAATGTTGTGACTTGGCGCCCGGAGAGCAGAGCCATTAAGTCTCTCAAGAGCGGC 6560
QY 1385 TTTATATGGGGGTTCCTGTACTAATTTCAAAAGGCGAGATCTGGGCTATGCGGGTCC 1444
DB 6561 TTTATATGGGGGTTCCTGTACTAATTTCAAAAGGCGAGATCTGGGCTATGCGGGTCC 6620
QY 1445 GCGGAGGCGGCTGTGACGACTAGCTGCGGTAATTAACCTCAGATGTTACTTGAAGGCT 1504
DB 6621 GCGGAGGCGGCTGTGACGAGCTAGCTGCGGTAATTAACCTCAGATGTTACTTGAAGGCT 6680
QY 1505 CTGAGCCTGTGAGCTGCAAGCTCCAGAGCTCAGAGCATGCTGTGTGAGAGAGAG 1564
DB 6681 CTGAGCCTGTGAGCTGCAAGCTCCAGAGCTCAGAGCATGCTGTGTGTGAGAGAGAG 6740
QY 1565 TTGTGCTTATCTGTGAGAGCGCGGAAACCAAGAGAGAGCGGCGAGGCTTACGAGTCTTCA 1624
DB 6741 TTGTGCTTATCTGTGAGAGCGCGGGAACCAAGAGAGAGCGGCGAGGCTTACGAGTCTTCA 6800
QY 1625 GCGAGGCTATGACTAAGGACTCTGCCCCCCCCCGGAGAACCGGCCCAACAGAAATACAGC 1684
DB 6801 GCGAGGCTATGACTAAGGACTCTGCCCCCCCCCGGAGAACCGGCCCAACAGAAATACAGC 6860
QY 1685 TGGAGTGTATTAACATCATGCTCTCTCAATGTGTGCTGCGCAGATGCTATGGCAAAA 1744
DB 6861 TGGAGTGTATTAACATCATGCTCTCTCAATGTGTGCTGCGCAGATGCTATGGCAAAA 6920
QY 1745 GGGTATTAATCACTCAACCGGTGACC 1769
DB 6921 GGGTATTAATCACTCAACCGGTGACC 6945
```

RESULT 8

AAD25329

ID AAD25329 standard; cDNA; 7987 BP.

XX



AC AAD25329;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Hepatitis C virus (HCV) adaptive replicon VII cDNA.  
XX  
KM Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
KM gene therapy; vaccine; immunoprotection; hepatotropic; viraemic; liver;  
KM adaptive replicon VII; ss.  
XX  
OS Hepatitis C virus.  
XX  
PN WO200189364-A2.  
XX  
PD 29-NOV-2001.  
XX  
PF 23-MAY-2001; 2001WO-US016822.  
XX  
PR 23-MAY-2000; 2000US-00576989.  
XX  
PA (UNIV ) UNIV WASHINGTON.  
XX  
PI Rice CM, Blight KJ;  
XX  
DR MPI; 2002-066755/09.  
XX  
PT Hepatitis C virus variants having greater transfection efficiency and  
ability to survive subpassage, useful as a vaccine for immunizing primate  
to the virus, comprise non-naturally occurring viral sequences.  
XX  
PS Disclosure; Page 84-87; 174pp; English.  
XX  
CC The invention relates to Hepatitis C virus (HCV) variants which include  
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
CC variants that have a transfection efficiency and ability to survive  
CC subpassage greater than HCV that have wild-type polypeptide coding  
CC regions. The polynucleotides of the invention are useful for identifying  
CC a cell line that is permissive for infection with HCV and detecting  
CC replication of HCV in cells of the cell line. They are also useful for  
CC testing a compound for anti-viral properties and for inhibiting HCV  
CC infection. They are also useful for the generation of defined HCV virus  
CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
CC attachment, penetration and entry, structure/function studies on HCV  
CC proteins and RNA elements and identification of new antiviral targets, a  
CC systematic survey of cell culture systems and conditions to identify  
CC those that support wild-type and variant HCV RNA replication and particle  
CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of attenuated or  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilisation of the HCV  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
CC The present sequence is Hepatitis C virus (HCV) adaptive replicon VII  
CC cDNA  
XX  
SQ Sequence 7987 BP; 1647 A; 2368 C; 2242 G; 1730 T; 0 U; 0 Other;  
Query Match 87.7%; Score 1552.2; DB 6; Length 7987;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;  
QY 5 ATGACTTCCAGAGCGTGAAGTCAATGAGGCAACCTCCGTGCGCGCATGAAATGGCG 64  
DB 5381 ATGACTTCCAGAGCGTGAAGTCAATGAGGCAACCTCCGTGCGCGCATGAAATGGCG 5440  
QY 65 GGGACATTAACCGCGTGAAGTCAAGAAACAAGTAGTAATCTGTGACCTTTTGACCCGCG 124

DB 5441 GGAACATCACCCGCGTGAAGTCAAGAAATTAAGTAGTAATTTTGAAGCTTTTGAGCCGCG 5500  
QY 125 TCCGAGCGAGAGAGATGAGCGGGAAGTGTCCGTCCCGGAGATCTTCGGAATCCA 184  
DB 5501 TCCAGCGAGAGAGATGAGAGGAAGTATCCGTTCGCGGAGATCTTCGGAAGTCCA 5560  
QY 185 AGAAATTTCCACAGAGAGATGCCGCAATGGGACAGCCCGGATTCACACCTCCGTGTGG 244  
DB 5561 GGAATTTCTCTGAGAGATGCCAATATGGGACAGCCCGGATTCACACCTCTTGA 5620  
QY 245 AGTCTGGAAGAGCCCGGAGATCAAGTCTCCAGTGTATCAATGAGTCCCATCTCA 304  
DB 5621 AGTCTGGAAGAGCCCGGAGATCAAGTCTCCAGTGTATCAATGAGTCCCATCTCA 5680  
QY 305 CTAAAGACCCCTCTTATACCACTTCACAGAGAAAGAGACAGTGTTCGACGAATCCA 364  
DB 5681 CCAAGGCCCTCCGAATACCACTTCACAGAGAAAGAGACAGTGTTCGACGAATCCA 5740  
QY 365 CCGTGTCTTCTGCGCCTTGAGGAGCTTGCCACAAAGGCTTTGTAGTCCGACCGTGG 424  
DB 5741 CCGTGTCTTCTGCGCCTTGAGGAGCTTGCCACAAAGGCTTTGTAGTCCGACCGTGG 5800  
QY 425 CCGTGCACAGCGGACAGGCAACCGCCCTCTGACCAATCTTCGACAGAGCGGAGAG 484  
DB 5801 CCGTGCACAGCGGACAGGCAACCGCCCTCTGACCAAGCTTCGACAGAGCGGAGAG 5860  
QY 485 GATCTGACGTTGATGTTGATTTCTTCATGTCCTCCCTTGAAGGGGAGCGGAGAGCCCG 544  
DB 5861 GATCCACAGTGTGATGTTGATTTCTTCATGTCCTCCCTTGAAGGGGAGCGGAGAGCCCG 5920  
QY 545 ATTCGACGAGCGGAGTCTGATCTACCGTAGTGAAGAGCGGAGAGAGCGTGTGCTGCT 604  
DB 5921 ATTCGACGAGCGGAGTCTGATCTACCGTAGTGAAGAGCGGAGAGAGCGTGTGCTGCT 5980  
QY 605 GCTGATGTCCTTACATGATGAGACAGGCGCTGTGATCAAGCATGCGTGGAGAGAAAGCA 664  
DB 5981 GCTGATGTCCTTACATGATGAGACAGGCGCTGTGATCAAGCATGCGTGGAGAGAAAGCA 6640  
QY 665 AGCTGCCATCAACGCGTTGAGCACTTTTGTGCTGCTGATCAACAAATGTGTACGCTA 724  
DB 6041 AGCTGCCATCAACGCGTTGAGCACTTTTGTGCTGCTGATCAACAAATGTGTACGCTA 6100  
QY 725 CCAATCCCGGAGCGGAGCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784  
DB 6101 CCAATCCCGGAGCGGAGCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6160  
QY 785 TGAGCATCACTACAG 844  
DB 6161 TGAGCATCACTACAG 6220  
QY 845 CTAAAGCTTATATGATGAG 904  
DB 6221 CTAAAGCTTATATGATGAG 6280  
QY 905 AATTTGGCATGAGGAG 964  
DB 6281 AATTTGGCATGAGGAG 6340  
QY 965 GCTCCGTTGAG 1024  
DB 6341 GCTCCGTTGAG 6400  
QY 1025 CAAAAAATAGAGTTTCTGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084  
DB 6401 CAAAAAATAGAGTTTCTGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6460  
QY 1085 TGGTGTTCACAGCTTGGGGGTCCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1144  
DB 6461 TGGTGTTCACAGCTTGGGGGTCCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6520  
QY 1145 CCAACCTCCCTGAG 1204  
DB 6521 CCAACCTCCCTGAG 6580

```
QY 1205 GGGTCAGTTCCTGGTGAAGCGCTGGAAATCAAGAGACCCCTATGGGCTTTGCATATG 1264
    |||||
DB 6581 GGGTCAGTTCCTGGTGAATGCTGGAAAGCAAGAAATGCCCTATGGGCTTCGATATG 6640
QY 1265 ACAACCCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCAATTT 1324
    |||||
DB 6641 ACAACCCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCAATTT 6700
QY 1325 ATCAATGTTGACTGTTGGCCCCCGAAGCAGACGGCCATAAGTGTGCTCAAGAGCGGC 1384
    |||||
DB 6701 ACCAATGTTGACTGTTGGCCCCCGAAGCAGACGGCCATAAGTGTGCTCAAGAGCGGC 6760
QY 1385 TTTATATCGGGGGGCTCCCTGACTMAATTCAAAAGGAGAACTGCGGCTATCCCGGTGC 1444
    |||||
DB 6761 TTTATATCGGGGGGCTCCCTGACTMAATTCAAAAGGAGAACTGCGGCTATCCCGGTGC 6820
QY 1445 GCGGAGCGCGGCTGCTGACGACTAGTGGGTAAATCCCTCAATGTTACTTGAAGGCTT 1504
    |||||
DB 6821 GCGGAGCGCGGCTGCTGACGACTAGTGGGTAAATCCCTCAATGTTACTTGAAGGCTT 6880
QY 1505 CTGACGCGCTGTGAGCTGCAAGCTCCAGACGACGATGCTGCTGTGCGGAGACGGCC 1564
    |||||
DB 6881 CTGACGCGCTGTGAGCTGCAAGCTCCAGACGACGATGCTGCTGTGCGGAGACGGCC 6940
QY 1565 TTGTGCTTATCTGTGAGAGCGCGGAAACCAAGAGAGACGCGGCTACGAGTCTTCA 1624
    |||||
DB 6941 TTGTGCTTATCTGTGAGAGCGCGGAAACCAAGAGAGACGCGGCTACGAGTCTTCA 7000
QY 1625 CGGAGGCTTATGACTAGTACTCTGCCCCCGGGGAGCCGCCCAACCAAGAAATGAGCC 1684
    |||||
DB 7001 CGGAGGCTTATGACTAGTACTCTGCCCCCGGGGAGCCGCCCAACCAAGAAATGAGCC 7060
QY 1685 TGGAGTGTAATCATCATGCTCTCCAAATGTTGCTGCGGAGACGATGCTGGCAAA 1744
    |||||
DB 7061 TGGAGTGTAATCATCATGCTCTCCAAATGTTGCTGCGGAGACGATGCTGGCAAA 7120
QY 1745 GGGTATACTACTCAACCCGCTGAGCC 1769
    |||||
DB 7121 GGGTATACTACTCAACCCGCTGAGCC 7145

RESULT 9
AAD25321 standard; cDNA; 7987 BP.
AAD25321;
AAD25321;
12-MAR-2002 (first entry)
Hepatitis C virus (HCV) repBartman/delta2U's cDNA.
Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
85.
XX
OS Hepatitis C virus.
XX
PN MO200189364-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US016822.
XX
PR 23-MAY-2000; 2000US-00576989.
XX
PA (UNITW ) UNIV WASHINGTON.
XX
PI Rice CM, Blight KJ;
XX
DR WPI; 2002-066755/09.
XX
PT Hepatitis C virus variants having greater transfection efficiency and
```

```
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
PS Claim 44; Page 66-69; 174pp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the liver
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) repBartman/delta2U's
CC cDNA
SQ
Sequence 7987 BP; 1647 A; 2368 C; 2243 G; 1729 T; 0 U; 0 Other;
87.7%; Score 1552.2; DB 6; Length 7987;
Query Match 92.5%; Pred. No. 0;
Beet Local Similarity 0; Mismatches 133; Indels 0; Gaps 0;
Matches 1632; Conservative 0;

QY 5 ATGACTTCCGAGAGGCTTAACCTCATCGAGGCCAACCTTCTGTGCGGATGATGGGCG 64
    |||||
DB 5381 ATGACTTCCGAGAGGCTTAACCTCATCGAGGCCAACCTTCTGTGCGGATGATGGGCG 5440
QY 65 GGGACATTACCCTGCGTGAAGTCAAGAAACAAGTATCTGAGACTCTTTCAGACCGCG 124
    |||||
DB 5441 GGGACATTACCCTGCGTGAAGTCAAGAAACAAGTATCTGAGACTCTTTCAGACCGCG 5500
QY 125 TCCGAGCGGAGAGATGAGCGGAAAGTTCCTCCGCGGAGATCTGCGGAATCCA 184
    |||||
DB 5501 TCCGAGCGGAGAGATGAGCGGAAAGTTCCTCCGCGGAGATCTGCGGAATCCA 5560
QY 185 AGAATTTCCACACAGCATGCCCCGATGAGGACCGCGGATTAACAACCTTCGCTGCTGG 244
    |||||
DB 5561 AGAATTTCCACACAGCATGCCCCGATGAGGACCGCGGATTAACAACCTTCGCTGCTGG 5620
QY 245 AGTCTGGAAGAGCCCGGACTACGTCCTCCAGTGTATCATGGGTGCCCATGCACTCA 304
    |||||
DB 5621 AGTCTGGAAGAGCCCGGACTACGTCCTCCAGTGTATCATGGGTGCCCATGCACTCA 5680
QY 305 CTAAGACCCCTCTAATACACCTCCACGAGAAAGAGACAGTTGTTTGAACAGATCCA 364
    |||||
DB 5681 CTAAGACCCCTCTAATACACCTCCACGAGAAAGAGACAGTTGTTTGAACAGATCCA 5740
QY 365 CCGTGTCTTTCGCTGCGGAGAGTTCGCAAAAGCTTTTGTGATGCTCCGAGACCTGTGG 424
    |||||
DB 5741 CCGTGTCTTTCGCTGCGGAGAGTTCGCAAAAGCTTTTGTGATGCTCCGAGACCTGTGG 5800
QY 425 CCGTGAAGGAGGAGCAACCGCCCTCTCGAACATCTCTCGAGAGACGCGGAGAG 484
    |||||
DB 5801 CCGTGAAGGAGGAGCAACCGCCCTCTCGAACATCTCTCGAGAGACGCGGAGAG 5860
QY 485 GATCTGACGTTGAGTGTATTCCTTCATGCCCCCTTGAAGGAGGAGCCGCGGAGACCCG 544
```



```

Db      5861 GATCCGACGTTGAGTGTACTCTCTCATGCCCCCTTGAGGGGAGCCGGGGATCCCG 5920
Qy      545 ATCTCAGCGAGCGGCTTGGTCTTACCGTGAATGAGAGCCGGTGAAGAGCTGCTGCT 604
Db      5921 ATCTCAGCGAGCGGCTTGGTCTTACCGTGAAGAGAGGCGTGAAGAGCTGCTGCT 5980
Qy      605 GCTCGAGTGTCTTACATAGAGAGGCGCTGTGATCAAGCCATGCGTGGAGAGAAAGCA 664
Db      5981 GCTCGAGTGTCTTACATAGAGAGGCGCGCTGTATCAAGCCATGCGTGGAGAGAAAGCA 6040
Qy      665 AGCTGCCATCAACCGGCTTGAAGCACTTTGCTGCTCAACCAATGCTTACGCTA 724
Db      6041 AGCTGCCATCAATGAGACTGAGCACTTTGCTGCTCAACCAATGCTTACGCTA 6100
Qy      725 CCAATATCCCGCAGCGGAGCGGAGGAGAGAGGCTTGTGACAGCTGCAAAATCC 784
Db      6101 CAACATCTCGCAGCGGAGCGGAGGAGAGAGGCTTGTGACAGCTGCAAGTCC 6160
Qy      785 TGGACGATCACTTCCAGAGAGCTGCTCAAGAGATGAAGGCGGAGGCTTCAAGTAAAG 844
Db      6161 TGGACGACCACTTCCAGAGAGCTGCTCAAGAGATGAAGGCGGAGGCTTCAAGTAAAG 6220
Qy      845 CTAAAGCTTCTATGATGAGAGAGGCTGCAAGGCTGACGCGCCACATTCGGCCAAATCTA 904
Db      6221 CTAAAGCTTCTATGATGAGAGAGGCTGCTGAGAGAGGCTGAGAGGCTTCAAGTAAAG 6280
Qy      905 AATTGGCTATGAGGAGCAAGAGAGCTGCGGAACTTATCCAGCAAGGCGCATTAACCATATCC 964
Db      6281 AATTGGCTATGAGGAGCAAGAGAGCTGCGGAACTTATCCAGCAAGGCGCATTAACCATATCC 6340
Qy      965 GCTCCGCTGAGGAGAGCTTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCT 1024
Db      6341 GCTCCGCTGAGGAGAGCTTGTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCT 6400
Qy      1025 CAAAAAATGAGGTTTTCTGCGTCAACAGAGAGAGAGGCGGCAAGGCGCTGCGCTTA 1084
Db      6401 CAAAAAATGAGGTTTTCTGCGTCAACAGAGAGAGAGGCGGCGGCGGCGGCGGCGGCTTA 6460
Qy      1085 TGTGTGTCCAGAGCTTGGGGGTCGTTGTGAGAGAGAGAGGCGGCTTATGAGAGTGTCT 1144
Db      6461 TGTGTGTCCAGAGCTTGGGGGTCGTTGTGAGAGAGAGAGGCGGCTTATGAGAGTGTCT 6520
Qy      1145 CCAACCTCCCTGAGGCTGTGATGAGGCTCCCTGTAAGGATTCAGATTTCTCTGAGAGAGC 1204
Db      6521 CCAACCTCCCTGAGGCTGTGATGAGGCTCCCTGTAAGGATTCAGATTTCTCTGAGAGAGC 6580
Qy      1205 GGGTCGAGTTCTGCTGTAAGAGCGCTGGAATCAAGAGAGAGAGGCGGCTTATGAGAGTATG 1264
Db      6581 GGGTCGAGTTCTGCTGTAAGAGCGCTGGAATCAAGAGAGAGAGGCGGCTTATGAGAGTATG 6640
Qy      1265 ACAACCGCTGTTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATTT 1324
Db      6641 ACAACCGCTGTTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATTT 6700
Qy      1325 ATCAATGTTGTGACTTGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1384
Db      6701 ACCAAATGTTGTGACTTGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6760
Qy      1385 TTTAATGCGGGGGTCCCTGACTAATTTCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1444
Db      6761 TTTAATGCGGGGGTCCCTGACTAATTTCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 6820
Qy      1445 GGGCGAGCGGCGTGTGACGACTAGTGGGTAAATACCTTCACATGTTACTTGAAGGCTT 1504
Db      6821 GGGCGAGCGGCGTGTGACGACTAGTGGGTAAATACCTTCACATGTTACTTGAAGGCTT 6880
Qy      1505 CTGACAGCTGTGAGAGTGCAGAAAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1564
Db      6881 CTGCGGCTGTGAGAGTGCAGAAAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6940
Qy      1565 TTGCTGTTATCTGTGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCTTCA 1624

```

```

Db      6941 TTGCTGTATCTGTGAAAGCGGGGAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7000
Qy      1625 CGAGAGCTATGACTAGTACTCTGCCCCCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1684
Db      7001 CGAGAGCTATGACTAGTACTCTGCCCCCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7060
Qy      1685 TGGAGTTGATTAACATGATGCTCTCCATATGTGCGGTGCGGAGAGAGAGAGAGAGAGAGAGAG 1744
Db      7061 TGGAGTTGATTAACATGATGCTCTCCATATGTGCGGTGCGGAGAGAGAGAGAGAGAGAGAG 7120
Qy      1745 GGGTATATCACTACCTACCCGCTGAGCC 1769
Db      7121 GGGTATATCACTACCTACCCGCTGAGCC 7145

RESULT 10
AAD25324
ID AAD25324 standard; cDNA; 7987 BP.
XX
AC AAD25324;
XX
DT 12-MAR-2002 (first entry)
XX
DE Hepatitis C virus (HCV) adaptive replicon VI cDNA mutant.
XX
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KW adaptive replicon VI; mutant; ss.
OS Hepatitis C virus.
OS Synthetic.
XX
FH Key
FH Location/Qualifiers
FT CDS
FT 1601..7758
FT /*tag= a
FT /product= "HCV adaptive replicon VI protein"
FT 4642..5982
FT /*tag= b
FT /product= "NS5A protein of HCV adaptive replicon VII"
FT /note= "CDS does not include both start and stop codon"
FT /partial
FT replace(5336, G)
FT mutation
XX
EN WO200189364-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US016822.
XX
PR 23-MAY-2000; 2000US-00576989.
XX
PA (UNITW ) UNIV WASHINGTON.
XX
PI Rice CM, Blight RJ;
XX
DR WPI; 2002-066755/09.
DR P-PSDB; AAB15720, AAB15721.
XX
PT Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
XX
PS disclosure; Page 74-77; 174dp; English.
XX
CC The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV

```

infection. They are also useful for the generation of defined HCV virus stocks to develop in vitro and in vivo assays for virus neutralisation, attachment, penetration and entry, structure/function studies on HCV proteins and RNA elements and identification of new antiviral targets, a systematic survey of cell culture systems and conditions to identify those that support wild-type and variant HCV RNA replication and particle release, production of adaptive HCV variants capable of more efficiency replication in cell culture, production of HCV variants with altered tissue or species tropism, establishment of alternative animal models for inhibitor evaluation including those supporting HCV variant replication, development of cell-free HCV replication assays, production of immunogenic HCV particles for vaccination, engineering of attenuated HCV derivatives as possible vaccine candidates, engineering of attenuated or defective HCV derivatives for expression of heterologous gene products for gene therapy and vaccine applications and for utilisation of the HCV glycoproteins for targeted delivery of therapeutic agents to the liver or other cell types with appropriate receptors. Vaccine comprising these sequences is useful for inducing immunoprotection to HCV in a primate. The present sequence is Hepatitis C virus (HCV) adaptive replicon VI cDNA mutant. This sequence is generated by the mutation g to t at position 5336 of HCVrep1bBartman/Availi cDNA

Sequence 7987 BP, 1646 A, 2368 C, 2243 G, 1730 T, 0 U, 0 Other;

Query Match 87.7%; Score 1552.2; DB 6; Length 7987;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```
QY 5 ATGACTTCCCAACGCTGACCTCATCGAGGCCAACCTCTGTGGGGGATGAGTGGGCG 64
DB 5381 ATGACTCCCGGACCGCTGACCTCATCGAGGCCAACCTCTGTGGGGGATGAGTGGGCG 5440
QY 65 GGGACATTACCCGCTGAGTCAGAGAAACAAGTAGTATCTGTGACTCTTTCGACCCGCG 124
DB 5441 GGAACATCACCCGCTGAGTCAGAGAAATAAGTAGTATTTTGGACTCTTTCGAGCCGCG 5500
QY 125 TCCGACGGAGAGATGAGCGGAAAGTGTCCGTCGGCGGAGATCTTCGCGAAATCCA 184
DB 5501 TCCAGACGGAGAGATGAGGAGAAAGTATCCGTTCCGGCGGAGATCTTCGCGAGTCCA 5560
QY 185 AGAATTTCCACACGATGATGCGGATGGGACGCGCGGATTAACAACCTTCGCTGCTGG 244
DB 5561 GGAATTTCCCTGACGATGCCATATGCGGACGCGCGGATTAACAACCTTCGCTGCTGG 5620
QY 245 AGTCTCGAAGGCGCCGAGCTACGTCCTCCAGTGGTATCAGGGTCCCACTGCAACCTA 304
DB 5621 AGTCTCGAAGGACCGGAGCTACGTCCTCCAGTGGTACGAGGGTCCCACTGCAACCTA 5680
QY 305 CTAAAGACCCCTCTATACCACTCCACGGAAGAGACAGTTGTTCTGACAGATCCA 364
DB 5681 CCAAGGCCCTCCGATACCACTCCACGGAAGAGAGCGGTTGTCCTGTCAGATCTTA 5740
QY 365 CCGTCTTTCGCGCTGCGGAGCTTGGCAAAAGCTTTTGTAGCTCCGACCGTGGG 424
DB 5741 CCGGTCTTTCGCGCTGCGGAGCTTGGCAAAAGCTTTGCGAGCTCCGATGCTGGG 5800
QY 425 CCGTGAACGCGGACGCGCAACGCGCCCTCTGACCAATCTCCGACGACGCGAGGAG 484
DB 5801 CCGTGAACGCGGACGCGCAACGCGCTCTCTGACCAACGCGCTCCGACGACGCGG 5860
QY 485 GATCTGACGTTGATGCTATTCCTCCAGGCGCGCCCTTGAAGGGAGAGCGCGGAGAC 544
DB 5861 GATCTGACGTTGATGCTATTCCTCCAGGCGCGCCCTTGAAGGGAGAGCGCGGAGAC 5920
QY 545 ATCTAGAGGACGGGCTTGGTCTACCGGTAGTGAAGAGCGCGTGAAGACGTCGTCGT 604
DB 5921 ATCTAGAGGACGGGCTTGGTCTACCGGTAGTGAAGAGCGCGTGAAGACGTCGTCGT 5980
QY 605 GCTCGATGCTTACACATGACAGGCGCTGTATCACGCGCATGCGCTGGCGAGGAAAGCA 664
DB 5981 GCTCGATGCTTACACATGACAGGCGCGCTGTATCACGCGCATGCGCTGGCGAGGAAAGCA 6040
QY 665 AGCTGCCATCAACGCGTTGAGCAACTCTTTGCTGGTCAACCAACAATGCTACGCTTA 724
```

```
DB 6041 AGCTGCCATCAATGACATGAGCAACTCTTTGCTCCGTCACACAACTTGGTCTATGCTTA 6100
QY 725 CCAATCCCGGACGCGCAACGCGGCGAGAAAGTAGTACTCTTGAACGACTGCAAAATCC 784
DB 6101 CAACATTCGAGCGCAACGCTGCGCGAGAAAGTAGTACTCTTGAACGACTGCAAGTCC 6160
QY 785 TGAAGATCACTACAGGACGCTGCTCAAGAGATGAAGCGGACGTCACAGTTAAGG 844
DB 6161 TGAAGATCACTACCGGAGCTGCTCAAGAGATGAAGCGGACGTCACAGTTAAGG 6220
QY 845 CTAAGCTTCTATCACTAGAGAGAAAGCTTGAAGCTGACGCGCCCACTTGGCCAAATCTA 904
DB 6221 CTAAGCTTCTATCACTAGAGAGAAAGCTTGAAGCTGACGCGCCCACTTGGCCAGATCTTA 6280
QY 905 AATTGGCTATGGGCGCAAGAGACGTCGCGAACCCTATCAAGAAAGCCATTAAACCAATCC 964
DB 6281 AATTGGCTATGGGCGCAAGAGACGTCGCGAACCCTATCAAGAAAGCCATTAAACCAATCC 6340
QY 965 GCTCCGTGGAGAGACTTGTGGAAGACATGAACACCAATTGACACACCATCATATGG 1024
DB 6341 GCTCCGTGGAGAGACTTGTGGAAGACATGAAGACATGAAGACCAATGACACCATCATATGG 6400
QY 1025 CAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGCGCGCAAGCCAGCTGCGCTTA 1084
DB 6401 CAAAAATGAGGTTTCTGCGTCCAAACGAGAGAGAGCGCGCAAGCCAGCTGCGCTTA 6460
QY 1085 TCGTCTCCGACACTGGGGGCTCGGTGTGTGCGGAGAAATGGCCCTCTATACAGGTCCT 1144
DB 6461 TCGTATTCAGAGATTTGGGGGTTGCTGTGTGCGGAGAAATGGCCCTTATACATGTGTCCT 6520
QY 1145 CCACCTCCCTCAGGCTGTGATGGGCTCTGTAGAGATTCAGATATCTCTCGAGACAGC 1204
DB 6521 CCACCTCCCTCAGGCGCGTATGGGCTCTGTAGAGATTCAGATATCTCTCGAGACAGC 6580
QY 1205 GGGTGAATCTCTGAGTGAACGCTTGAATCAAAAGAGACCCCTATGGGCTTTGCAATAG 1264
DB 6581 GGGTGAATCTCTGAGTGAATCTTGAATCAAAAGAGAGACCCCTATGGGCTTTGCAATAG 6640
QY 1265 ACACCCGCTGTTTGAACAGCAACGTCAGTGAATGACATCCGTTGAGAGAGTCAATTT 1324
DB 6641 ACACCCGCTGTTTGAACAGCAACGTCAGTGAATGACATCCGTTGAGAGAGTCAATCT 6700
QY 1325 ATCAATGTTGACTGAGCTGAGCCCGGAGAGCCAGCATTAAGTCTGTCAGAGCGGC 1384
DB 6701 ATCAATGTTGACTGAGCTGAGCCCGGAGAGCCAGCATTAAGTCTGTCAGAGCGGC 6760
QY 1385 TTTATATCGGGGTTCCCTGACTAATTCAAAAGGCGAGAACTGCGGCTATCGCGGTGCC 1444
DB 6761 TTTATATCGGGGTTCCCTGACTAATTCAAAAGGCGAGAACTGCGGCTATCGCGGTGCC 6820
QY 1445 GCGGAGGCGGTGCTGACAGCTAGCGCGGTTATPACCTTCACAGTTACTTTGAAGGCT 1504
DB 6821 GCGGAGGCGGTGCTGACAGCTAGCGCGGTTATPACCTTCACAGTTACTTTGAAGGCT 6880
QY 1505 CTGACGCTGTCGAGCTGCAAGCTCCAGAGCTGCAAGATGCTGATGCGGAGAGCGGC 1564
DB 6881 CTGACGCTGTCGAGCTGCAAGCTCCAGAGCTGCAAGATGCTGATGCGGAGAGAGCGGC 6940
QY 1565 TTGTCTGTTATGTGAGAGCGCGGAAACCAAGAGAGACGCGGAGAGCTTACGAGTCTTCA 1624
DB 6941 TTGTCTGTTATGTGAGAGCGCGGAAACCAAGAGAGACGCGGAGAGCTTACGAGTCTTCA 7000
QY 1625 CCGAGGCTATGACATGATCTCTGCGCGCGCGCGCGCGCGCGCGCAACCAATATGAGCC 1684
DB 7001 CCGAGGCTATGACATGATCTCTGCGCGCGCGCGCGCGCGCGCGCAACCAATATGAGCC 7060
QY 1685 TGAAGTTGATTAACATCATGCTCTCTCAATGTGTGCGTGCAGCATGTCATCTGGCAAAA 1744
DB 7061 TGAAGTTGATTAACATCATGCTCTCTCAATGTGTGCGTGCAGCATGTCATCTGGCAAAA 7120
QY 1745 GGGTATACTACCTGACCCGTTGAGCC 1769
```



```

Db      6641 ACACCCGCTGTTTGACTCAAGCGTCACTAGAAATGACATCCGTTTGAAGAGTCAATCT 6700
Qy      1325 ATCAATGTTGACTTGGCCCCCGAAGCGAGACAGGCGCTTAAGTGCCTCACAGAGCGGC 1384
Db      6701 ACCAATGTTGTACTTGGCCCCCGAAGCGAGACAGGCGCTTAAGTGCCTCACAGAGCGGC 6760
Qy      1385 TTTATATCGGGGGTCCCTGACTTAATTCAAAAGGCGAAGCTGCGGCTATCGCCGGTGC 1444
Db      6761 TTTACATCGGGGGCCCCCTGACTTAATTCAAAAGGCGAAGCTGCGGCTATCGCCGGTGC 6820
Qy      1445 GGGCGAGCGCGCTGCTGAGAGTACTGCGGTAAATACCTTCACATGTTACTTGAAGGCT 1504
Db      6821 GGGCGAGCGCGTGTACTGAGACAGCTGCGGTAAATACCTTCACATGTTACTTGAAGGCGC 6880
Qy      1505 CTGCAAGCTGTGAGCTGCAAGCTCCAGAGCTGCAAGTGTGCTGCTGTCGAGAGCGGC 1564
Db      6881 CTGCGGCTGTGAGCTGCAAGCTCCAGAGCTGCAAGTGTGCTGCTGTCGAGAGAGAC 6940
Qy      1565 TTGTCGTATCTGTGAGAGCGCGGAAACCCAGAGAGACGCGGCGAGCTTACAGTCTTCA 1624
Db      6941 TTGTCGTATCTGTGAAAGCGCGGGAACCAAGAGAGACGAGCGAGCTTACAGGCTTCA 7000
Qy      1625 CGGAGGCTATGACTAGTACTGTGCCCCCGGGGACCCGCCCAACTCAAGAAATACGAC 1684
Db      7001 CGGAGGCTATGACTAGTACTGTGCCCCCGGGGACCCGCCCAACTCAAGAAATACGACT 7060
Qy      1685 TGGAGTTGATACATCATCATCTCTCCCAATGTGTGCGTGCAGAGTGCATCGGCAAA 1744
Db      7061 TGGAGTTGATACATCATCTCTCTCCCAATGTGTGTGCGTGCAGAGTGCATCGGCAAA 7120
Qy      1745 GGGTACTACTACCTCAACCCGCTGACCC 1769
Db      7121 GGGTACTACTACTCAACCCGCTGACCC 7145

```

## RESULT 12

AAD25322 standard; cDNA; 7989 BP.

AAD25322;

12-MAR-2002. (first entry)

Hepatitis C virus (HCV) rep/Bartman/Availi cDNA.

Hepatitis C virus; HCV; transfection; infection; virus neutralisation;

gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;

86.

Hepatitis C virus.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

XX      Hepatitis C virus variants having greater transfection efficiency and
PT      ability to survive subpassage, useful as a vaccine for immunizing primate
PT      to the virus, comprise non-naturally occurring viral sequences.
PS      Claim 44; Page 69-71; 174p; English.
XX      The invention relates to Hepatitis C virus (HCV) variants which include
XX      polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC      variants that have a transfection efficiency and ability to survive
CC      subpassage greater than HCV that have wild-type polypeptide coding
CC      regions. The polynucleotides of the invention are useful for identifying
CC      a cell line that is permissive for infection with HCV and detecting
CC      replication of HCV in cells of the cell line. They are also useful for
CC      testing a compound for anti-viral properties and for inhibiting HCV
CC      infection. They are also useful for the generation of defined HCV virus
CC      stocks to develop in vitro and in vivo assays for virus neutralisation,
CC      attachment, penetration and entry, structure/function studies on HCV
CC      proteins and RNA elements and identification of new antiviral targets, a
CC      systematic survey of cell culture systems and conditions to identify
CC      those that support wild-type and variant HCV RNA replication and particle
CC      release, production of adaptive HCV variants capable of more efficiency
CC      replication in cell culture, production of HCV variants with altered
CC      immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC      derivatives as possible vaccine candidates, engineering of attenuated or
CC      defective HCV derivatives for expression of heterologous gene products
CC      for gene therapy and vaccine applications and for utilisation of the HCV
CC      glycoproteins for targeted delivery of therapeutic agents to the liver
CC      or other cell types with appropriate receptors. Vaccine comprising these
CC      sequences is useful for inducing immunoprotection to HCV in a primate.
CC      The present sequence is Hepatitis C virus (HCV) rep/Bartman/Availi cDNA
XX      SQ

```

Sequence 7989 BP; 1647 A; 2369 C; 2242 G; 1731 T; 0 U; 0 Other;

Query Match 87.7%; Score 1552.2; DB 6; Length 7989;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```

Qy      5 ATGACTTCCAGAGCGTGAACCTCATCGAGCGCAACCTCTGTGCGCGATGATGGCG 64
Db      5381 ATGACTTCCAGAGCGTGAACCTCATCGAGCGCAACCTCTGTGCGCGATGATGGCG 5440
Qy      65 GGGACATTAACCGGTGAGTCAAGAAAGTAAAGTAAATTTGACTCTTTCAGCGCC 124
Db      5441 GGAACATTAACCGGTGAGTCAAGAAAGTAAAGTAAATTTGACTCTTTCAGCGCC 5500
Qy      125 TCCGAGCGAGAGAGATGAGCGGAAAGTGTCCGTCGCGAGATCCTGCGGAATCCA 184
Db      5501 TCCGAGCGAGAGAGATGAGCGGAAAGTGTCCGTCGCGAGATCCTGCGGAATCCA 5560
Qy      185 AGAAATTCACACAGCGATGCGCGATGCGCGCGCGGATTAACAACCTCTCGCTGTGG 244
Db      5561 GGAATTCCTCGAGCGATGCGCGATGCGCGCGCGGATTAACAACCTCTCGCTGTGG 5620
Qy      245 AGTCTGGAAGCGCGCGAGTCAAGTGTCCGTCGCGAGTCAAGTGTCCGTCGCGAGTCA 304
Db      5621 AGTCTGGAAGCGCGCGAGTCAAGTGTCCGTCGCGAGTCAAGTGTCCGTCGCGAGTCA 5680
Qy      305 CTAAGACCCCTCTATACCACTCAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 364
Db      5681 CTAAGACCCCTCTATACCACTCAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5740
Qy      365 CCGTGTCTTCTGCGCTGCGAGCTTTCACAAAGGCTTTTGTAGCTCGGACGCTGG 424
Db      5741 CCGTGTCTTCTGCGCTGCGAGCTTTCACAAAGGCTTTTGTAGCTCGGACGCTGG 5800
Qy      425 CCGTGTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484
Db      5801 CCGTGTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5860

```

QY 485 GATCTGACGTTGAGTGTCTCTCCATGAGCCCCCTTGAGGGGGAGCCGGGGAGCCCG 544  
 DB 5861 GATCCAGCTTGAAGTGTACTCTCCATGAGCCCCCTTGAGGGGGAGCCGGGGAGCTCCG 5920  
 QY 545 ATCTCAGCAGCGGGCTTGGTCTACCGTGAAGAGGCGGGTGAAGAGCTGTGTCT 604  
 DB 5921 ATCTCAGCAGCGGGCTTGGTCTACCGTGAAGAGGCTGAAGAGAGTGTGTCTGTCT 5980  
 QY 605 GCTCCATGCTCTACATGAGCAGAGGCTCTGTATCAACGCTATCCGCTGGGAGAGAAAGCA 664  
 DB 5981 GCTCGATGCTCTACATGAGCAGAGGCTCTGTATCAACGCTATCCGCTGGGAGAGAAAGCA 6040  
 QY 665 AGCTGCCCATCAACGCGTTGAGCACTCTTGTGTGTCACCAATGCTGTACGCTA 724  
 DB 6041 AGCTGCCCATCAATGAGCACTGTGTCACCAATGCTGTACGCTGTATGCTA 6100  
 QY 725 CCACATCCCGCAGCGCAGCAGCGCAGAGAGAGAGTCACTTTGACAGACTGCAATCC 784  
 DB 6101 CAACATCTCGACGCGCAGCAGCGCAGAGAGAGTCACTTTGACAGACTGCAATCC 6160  
 QY 785 TGAAGCATCACTACAGAGAGCTGTCAAGAGAGTGAAGGCGGCTCCAGTTAAG 844  
 DB 6161 TGAAGCATCACTACAGAGAGCTGTCAAGAGAGTGAAGGCGGCTCCAGTTAAG 6220  
 QY 845 CTAAAGCTTCTATCAGTGAAGAGAGCTGAAGCTGAGCGCCCACTGGGCCCAATCTA 904  
 DB 6221 CTAAAGCTTCTATCAGTGAAGAGAGCTGTGAGAGCTGAGCGCCCACTGGGCCCAATCTA 6280  
 QY 905 AATTGGCTATGAGGGGCAAGAGAGCTGCGGAACCTATTCAGCAGAGCCATTAACCAATCC 964  
 DB 6281 AATTGGCTATGAGGGGCAAGAGAGCTGCGGAACCTATTCAGCAGAGCCATTAACCAATCC 6340  
 QY 965 GCTCCGTTGGAGAGACTTGTGGAAGACACTGAACACCAATTAACCAATCAATGAG 1024  
 DB 6341 GCTCCGTTGGAGAGACTTGTGGAAGACACTGAAGACCAATTAACCAATCAATGAG 6400  
 QY 1025 CAAAAAATAGAGTTTCTCGTCCACAGAGAGAGGCGGCAAGCAGCTCGCTTA 1084  
 DB 6401 CAAAAAATAGAGTTTCTCGTCCACAGAGAGAGGCGGCAAGCAGCTCGCTTA 6460  
 QY 1085 TCGTGTTCACAGACTTGGGGGCTCCGTGTGTCAGAGAAATGGCCCTATGAGAGTGTCT 1144  
 DB 6461 TCGTGTTCACAGACTTGGGGGCTCCGTGTGTCAGAGAAATGGCCCTATGAGAGTGTCT 6520  
 QY 1145 CCAACCTCCCTCAGAGCTGTGATGGGCTCTCGTACGATTCAGATTCCTCTGAGACAG 1204  
 DB 6521 CCAACCTCCCTCAGAGCTGTGATGGGCTCTCGTACGATTCAGATTCCTCTGAGACAG 6580  
 QY 1205 GGGTGGAGTTCCTGGTGAAGCGCTGGAATCAAGAGAGCCCTAATGGGCTTGCATATG 1264  
 DB 6581 GGGTGGAGTTCCTGGTGAAGCGCTGGAATCAAGAGAGCCCTAATGGGCTTGCATATG 6640  
 QY 1265 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTGAGAGAGTCAATTT 1324  
 DB 6641 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTGAGAGAGTCAATTT 6700  
 QY 1325 ATCAATGTTGTGACTTGGCCCCCGAAGCAGACAGGCGCATAGGTGCTTCAAGAGCGGC 1384  
 DB 6701 ACCAATGTTGTGACTTGGCCCCCGAAGCAGACAGGCGCATAGGTGCTTCAAGAGCGGC 6760  
 QY 1385 TTTTATATGGGGGGTCCCTGACCTAATTTCAAAAGGGAGAACTGGGCTATCGCGGTGCC 1444  
 DB 6761 TTTTATATGGGGGGTCCCTGACCTAATTTCAAAAGGGAGAACTGGGCTATCGCGGTGCC 6820  
 QY 1445 GGGCGAGCGGCTGTGCTGACGACTAGCTGGGTAAATACCTTCAATTTTGAAGGCTT 1504  
 DB 6821 GGGCGAGCGGCTGTGCTGACGACTAGCTGGGTAAATACCTTCAATTTTGAAGGCTT 6880  
 QY 1505 CTGACAGCTGTGAGAGCTGCAAAAGCTCAAGAGCTGCAAGTCTGTGTGGGAGAGCGCC 1564  
 DB 6881 CTGACAGCTGTGAGAGCTGCAAAAGCTCAAGAGCTGCAAGTCTGTGTGGGAGAGCGCC 6940  
 QY 1565 TTGTGCTTATCTGTGAGAGCGGGAGAACCAAGAGAGACCGCGGACCTTACAGAGTCTTCA 1624

DB 6941 TTGTGCTTATCTGTGAAGCGGGAGACCCAGAGAGAGACAGAGGCTACCGGCTTCA 7000  
 QY 1625 CGAGGCTATGACTAGTACTTGTGCCCCCGGGAGACCCGCCCAACAGAAATAGACC 1684  
 DB 7001 CGAGGCTATGACTAGTACTTGTGCCCCCGGGAGACCCGCCCAACAGAAATAGAGCT 7060  
 QY 1685 TGAAGTGAATATCATGAGTCTTCCATGATCGGTGGGAGAGAGTGAATCTGGCAAAA 1744  
 DB 7061 TGAAGTGAATATCATGAGTCTTCCATGATGTCAGTGGGAGAGAGTGAATCTGGCAAAA 7120  
 QY 1745 GGGTATTAATCACTCAACCCGCTGACCC 1769  
 DB 7121 GGGTATTAATCACTCAACCCGCTGACCC 7145

RESULT 13  
 AAD25326  
 ID AAD25326 standard; cDNA; 7989 BP.  
 AC AAD25326;  
 DT 12-MAR-2002 (first entry)  
 DE Hepatitis C virus (HCV) adaptive replicon V mutant cDNA.  
 XX  
 XX Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
 KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
 KM adaptive replicon V; mutant; ss.  
 XX  
 OS Hepatitis C virus.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 4642..5982  
 FT /\*tag= a  
 FT /product= "NS5A variant of HCV adaptive replicon V"  
 FT /note= "CDS does not include both start and stop codon"  
 FT mutation  
 FT replace(5320, G)  
 FT /\*tag= b  
 XX  
 PN MO200189364-A2.  
 XX  
 PD 29-NOV-2001.  
 PD XX  
 PF 23-MAY-2001; 2001WO-US016822.  
 PF XX  
 PR 23-MAY-2000; 2000US-00576989.  
 PR XX  
 PA (UNITW ) UNIV WASHINGTON.  
 PA XX  
 PI Rice CM, Blight KJ;  
 PI XX  
 DR WPI: 2002-066755/09.  
 DR P-PSDB; AAE15724.  
 DR XX  
 PT Hepatitis C virus variants having greater transfection efficiency and  
 PT ability to survive subpassage, useful as a vaccine for immunizing primate  
 PT to the virus, comprise non-naturally occurring viral sequences.  
 PS  
 PS Disclosure; Page 80-83; 174pp; English.  
 XX  
 XX The invention relates to Hepatitis C virus (HCV) variants which include  
 CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
 CC variants that have a transfection efficiency and ability to survive  
 CC subpassage greater than HCV that have wild-type polypeptide coding  
 CC regions. The polynucleotides of the invention are useful for identifying  
 CC a cell line that is permissive for infection with HCV and detecting  
 CC replication of HCV in cells of the cell line. They are also useful for  
 CC testing a compound for anti-viral properties and for inhibiting HCV  
 CC infection. They are also useful for the generation of defined HCV virus  
 CC stocks to develop in vitro and in vivo assays for virus neutralisation,





RESULT 14  
AAD25325 standard; cDNA; 7989 BP.  
ID AAD25325; standard; cDNA; 7989 BP.  
AC AAD25325;  
XX  
XX 12-MAR-2002 (first entry)  
XX  
XX Hepatitis C virus (HCV) adaptive replicon II cDNA mutant.  
XX  
XX Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
XX gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
XX adaptive replicon II; mutant; ss.  
XX  
XX Hepatitis C virus.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
XX CDS 1801..7758  
XX /tag= a  
XX /product= "HCV adaptive replicon II protein"  
XX CDS 4642..5982  
XX /tag= b  
XX /product= "NS5A protein of HCV adaptive replicon II"  
XX /note= "CDS does not include both start and stop codon"  
XX /partial  
XX  
XX MO200189364A2.  
XX  
XX 29-NOV-2001.  
XX  
XX 23-MAY-2001; 2001MO-US016822.  
XX  
XX 23-MAY-2000; 2000US-00576989.  
XX  
XX (UNIM) UNIV WASHINGTON.  
XX  
XX Rice CM, Blight RJ;  
XX  
XX MPI; 2002-066755/09.  
XX P-PSDB; AAB15722, AAB15723.  
XX  
XX Hepatitis C virus variants having greater transfection efficiency and  
XX ability to survive subpassage, useful as a vaccine for immunizing primate  
XX to the virus, comprise non-naturally occurring viral sequences.  
XX  
XX Example 1; Page 77-80; 174pp; English.  
XX  
XX The invention relates to Hepatitis C virus (HCV) variants which include  
XX polynucleotides comprising non-naturally occurring HCV sequence and HCV  
XX variants that have a transfection efficiency and ability to survive  
XX subpassage greater than HCV that have wild-type polypeptide coding  
XX regions. The polynucleotides of the invention are useful for identifying  
XX a cell line that is permissive for infection with HCV and detecting  
XX replication of HCV in cells of the cell line. They are also useful for  
XX testing a compound for anti-viral properties and for inhibiting HCV  
XX infection. They are also useful for the generation of defined HCV virus  
XX stocks to develop in vitro and in vivo assays for virus neutralisation,  
XX attachment, penetration and entry, structure/function studies on HCV  
XX proteins and RNA elements and identification of new antiviral targets, a  
XX systematic survey of cell culture systems and conditions to identify  
XX those that support wild-type and variant HCV RNA replication and particle  
XX release, production of adaptive HCV variants capable of more efficiency  
XX replication in cell culture, production of HCV variants with altered  
XX tissue or species tropism, establishment of alternative animal models for  
XX inhibitor evaluation including those supporting HCV variant replication,  
XX development of cell-free HCV replication assays, production of  
XX immunogenic HCV particles for vaccination, engineering of attenuated HCV  
XX derivatives as possible vaccine candidates, engineering of attenuated or  
XX defective HCV derivatives for expression of heterologous gene products  
XX for gene therapy and vaccine applications and for utilisation of the HCV  
XX glycoproteins for targeted delivery of therapeutic agents to the liver

CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
CC The present sequence is Hepatitis C virus (HCV) adaptive replicon II cDNA  
CC mutant. This sequence is generated by the mutation a to g at position  
CC 5289 of HCVrep1bBartman/Avail cDNA  
XX  
XX  
SQ Sequence 7989 BP; 1644 A; 2369 C; 2245 G; 1731 T; 0 U; 0 Other;  
Query Match 87.7%; Score 1552.2; DB 6; Length 7989;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;  
5 ATGACTTCCGAGCGGTGACCTCATGAGGCGCAACCTCGTGGCGGAGTGAATGGGCG 64  
5381 ATGACTTCCGAGCGGTGACCTCATGAGGCGCAACCTCGTGGCGGAGTGAATGGGCG 5440  
65 GGGACATTAACCCGCGTGAATGAGAGCAAGAGTAACTCTGACTCTTTTCGACCGC 124  
5441 GGAACATTAACCCGCGTGAATGAGAGCAAGAGTAACTCTGACTCTTTTCGACCGC 5500  
125 TCCGAGCGGAGAGAGTGAAGCGGAAAGTTCCTCCGCGGAGATCTCGGAAATCCA 184  
5501 TCCGAGCGGAGAGAGTGAAGCGGAAAGTTCCTCCGCGGAGATCTCGGAGATCCA 5560  
185 AGAAATTCACACGAGGATGCGCATGGGCAAGCCCGATTACACCTCCGCTGCTGG 244  
5561 GGAATTCACACGAGGATGCGCATGGGCAAGCCCGATTACACCTCCGCTGCTGG 5620  
245 AGTCTGGAAGGCGCGGAGTGAAGTCTCAAGTGTGATGAGTCCGCTCCACTCACTA 304  
5621 AGTCTGGAAGGCGCGGAGTGAAGTCTCAAGTGTGATGAGTCCGCTCCACTCACTA 5680  
305 CTAAAGACCCCTCTTATACCACTCCAGAGAAAGAGACAGTTGTTCTGACAAATCCA 364  
5681 CCAAGGCGCGCTCGATACCACTCCAGAGAAAGAGACAGTTGTTCTGACAAATCCA 5740  
365 CGGTGCTTCTGCGCGGAGCTTGCAAGAGCTTTGTTGTTGTTGTTGTTGTTGTTGTT 424  
5741 CGGTGCTTCTGCGCGGAGCTTGCAAGAGCTTTGTTGTTGTTGTTGTTGTTGTTGTT 5800  
425 CCGTCGACAGCGGACGAGCAAGCGCCCTCTGACCAATCTCTCGACGAGCGGACAG 484  
5801 CCGTCGACAGCGGACGAGCAAGCGCCCTCTGACCAATCTCTCGACGAGCGGACAG 5860  
485 GATCTGACGTTGATGATGATATCTCTCATGCCCCCTTGAAGAGGAGCGGAGCCCG 544  
5861 GATCTGACGTTGATGATGATATCTCTCATGCCCCCTTGAAGAGGAGCGGAGCCCG 5920  
545 ATCTGACGACGCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 604  
5921 ATCTGACGACGCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5980  
605 GGTGATGTTCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664  
5981 GGTGATGTTCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6040  
665 AGTGCCTCAACAGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 724  
6041 AGTGCCTCAACAGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6100  
725 CCAATCTGACGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 784  
6101 CCAATCTGACGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 6160  
785 TGGACGATCACTACAGAGAGTGTCTCAAGAGATGAGAGCGGAGCGGAGCGGAGAGG 844  
6161 TGGACGATCACTACAGAGAGTGTCTCAAGAGATGAGAGCGGAGCGGAGAGGAGG 6220  
845 CTAAAGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 904  
6221 CTAAAGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6280  
905 AATTGGCTATGAGGAG 964

```
DB 6281 AATTGGCTATGGGGGAAAGAGAGTCGGAACCTATCCAGCAAGGCCGTTAACACATCC 6340
QY 965 GGTCCGTGGGAGAGAGCTTTGGAGAGAGACAGCAATTTGACACACATCANTAG 1024
DB 6341 GCTCCGTGGAGAGAGCTTTGGAGAGAGACAGCAATTTGACACACATCANTAG 6400
QY 1025 CAAAAAATGAGGTTTCTGCTGCAACAGAGAGAGAGGCGCAAGCAGCTGCGCTTA 1084
DB 6401 CAAAAAATGAGGTTTCTGCTGCAACAGAGAGAGAGGCGCGCAAGCAGCTGCGCTTA 6460
QY 1085 TGTGTTCCCAAGCTTTGGGGGTCCTGTGTGGAGAAAATGGCCCTTAATGACGTGCT 1144
DB 6461 TGTGTTCCCAAGCTTTGGGGGTCCTGTGTGGAGAAAATGGCCCTTAATGACGTGCT 6520
QY 1145 CCACCTCCCTCAGAGCTGATGGGCTCCTCGTACGAGATTCAGATTTCTCTGACAGC 1204
DB 6521 CCACCTCCCTCAGAGCTGATGGGCTCCTCGTACGAGATTCAGATTTCTCTGACAGC 6580
QY 1205 GGGTGAAGTTCCTGGTGAACGCTGGAATCAAGAGAGACCCCTATGGGCTTTGATATG 1264
DB 6581 GGGTGAAGTTCCTGGTGAATGCTGGAAGAGCAAGAAATGCCCTATGGGCTTCGATATG 6640
QY 1265 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATT 1324
DB 6641 ACACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGTAGAGAGTCAATT 6700
QY 1325 ATCAATGTTGACTGTTGGCCCCCGAAGCAGACGCGCATTAAGTCCGTCACAGAGCGC 1384
DB 6701 ACCAATGTTGACTGTTGGCCCCCGAAGCAGACGCGCATTAAGTCCGTCACAGAGCGC 6760
QY 1385 TTTATATCGGGGGGTCCTGACTAAATTCAAAAGGAGAGAGTGGGCTATCGCCGGTGC 1444
DB 6761 TTTATATCGGGGGGTCCTGACTAAATTCAAAAGGAGAGAGTGGGCTATCGCCGGTGC 6820
QY 1445 GCGCGAGCGGCTGCTGACGACTAGTGGGTAATCCCTCACATGTTACTTGAAGGCT 1504
DB 6821 GCGCGAGCGGCTGCTGACGACTAGTGGGTAATCCCTCACATGTTACTTGAAGGCT 6880
QY 1505 CTGACGCTGTCGAGCTGCAAGGCTCCAGAGCTGCAAGCTGCTCGTGGCGGAGCGGC 1564
DB 6881 CTGACGCTGTCGAGCTGCAAGGCTCCAGAGCTGCAAGCTGCTCGTGGCGGAGCGGC 6940
QY 1565 TTGTGTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACGATCTTCA 1624
DB 6941 TTGTGTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACGATCTTCA 7000
QY 1625 CGGAGGCTATGACTAGTACTCTGCCCCCGGGGAGCCGCCGCCAACAGAAATACGACT 1684
DB 7001 CGGAGGCTATGACTAGTACTCTGCCCCCGGGGAGCCGCCGCCAACAGAAATACGACT 7060
QY 1685 TGGAGTTGATTAATCATCATCTCTCCCAATGTGTCGGTGGCGAGCATGTCGCAAAA 1744
DB 7061 TGGAGTTGATTAATCATCATCTCTCCCAATGTGTCGGTGGCGAGCATGTCGCAAAA 7120
QY 1745 GGGTATTAATCACTCAACCCGTGACCC 1769
DB 7121 GGGTATTAATCACTCAACCCGTGACCC 7145

RESULT 15
ADD93722 standard; DNA; 7989 BP.
XX
AC ADD93722;
XX
DT 29-JAN-2004 (first entry)
XX
DE Hepatitis C virus strain H77 (BB7-F1) replicon.
XX
KM HCV; vaccine; virulence; ss.
XX
OS Hepatitis C virus.
```

```
XX
PN WO2003085084-A2.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2003; 2003WO-US010177.
XX
PR 03-APR-2002; 2002US-0369685P.
XX
PA (SMK) SMITHLINE BECHAM CORP.
XX
PI Gates A, Gu B, Sarisky RT;
XX
DR WPI; 2003-804301/75.
XX
PT New hepatitis C virus (HCV) sub-genomic replicon, useful for facilitating
PT screening or testing of anti-HCV drugs, comprises a nucleic acid
PT construct encoding chimeric HCV non-structural proteins, and an NS5B
PT polymerase gene.
XX
PS Claim 16; Page 27-32; 159pp; English.
XX
CC The present sequence comprises a replicating hepatitis C virus (HCV) H77
CC (BB7-F1) sub-genomic replicon. The invention provides sub-genomic
CC replicons of HCV comprising a nucleic acid construct encoding chimeric
CC HCV nonstructural protein and an NS5B polymerase gene. A preferred
CC replicon comprises an NS3 nucleotide sequence ADD93721 that encodes the
CC first 75 contiguous N-terminal amino acids of the NS3 of genotype 1b, of
CC a BB7 strain. A chimeric replicon may comprise an NS3 sequence from any
CC of the 6 major HCV genotypes and subtypes but has its first 225
CC nucleotides of the coding sequence replaced by the BB7 strain NS3
CC sequence, especially where the replicon is from HCV genotype 1a (H77
CC strain) or genotype 1b (J4 strain). Stable cell lines expressing and
CC replicating functional replicons containing sequences from HCV genotype
CC 1a (strain H77) or genotype 1b (strain J4) within the prototype 1b
CC replicon backbone from HCV strain BB7 are provided. These can be used to
CC screen for compounds that modulate viral replication. The sub-genomic HCV
CC replicon systems of the invention may provide the foundation for
CC generating HCV replicons of all 6 major genotypes and subtypes to
CC facilitate screening, testing and evaluating anti-infective agents for
CC HCV disease(s).
XX
SQ Sequence 7989 BP; 1626 A; 2373 C; 2235 G; 1755 T; 0 U; 0 Other;
XX
Query Match 87.7%; Score 1552.2; DB 10; Length 7989;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1032; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 5 ATGACTTCCAGACGCTGACCTCATCGAGGCCAACTCTGTGGCGGATGAGATGGGCG 64
DB 5381 ATGACTCCCGGAGCGCTGACCTCATCGAGGCCAACTCTGTGGCGGAGATGGGCG 5440
QY 65 GGGACATTACCCCGGTGAGTCAAGAGAACAGGTAATCTTGACTTTTCCACCCGC 124
DB 5441 GGAACATACCCCGGTGAGTCAAGAGAACAGGTAATTTTGGACTTTTCCAGCCGC 5500
QY 125 TCCGAGCGGAGAGATGAGCGGGAAGTCCGTCGCGGCGAGATCTTGCGGAATCA 184
DB 5501 TCCAGCGGAGAGATGAGCGGGAAGTCCGTCGCGGCGAGATCTTGCGGAGTCA 5560
QY 185 AGAATTTCCACAGCGATGCGCCGATGGGACGCCCGGATTAACAACCTTCGCTGCG 244
DB 5561 GGAATTTCCGTCGAGCGATGCGCCGATGGGACGCCCGGATTAACAACCTTCGCTGAG 5620
QY 245 AGTCTGGAAGGCGCCGAGCTAGTCCCTCCAGTGTGTCATGGGTGCCACTGCCACTA 304
DB 5621 AGTCTGGAAGGCGCCGAGCTAGTCCCTCCAGTGTGTCATGGGTGCCACTGCCACTG 5680
QY 305 CTAAGACCCCTCTATACCACTTCAAGAGAGAGAGAGAGAGAGTGTGTGACAGAAATCA 364
DB 5681 CTAAGACCCCTCTCTATACCACTTCAAGAGAGAGAGAGAGAGAGTGTGTGACAGAAATCA 5740
QY 365 CCGTGTCTTTCGCTGCGGAGCTTGCACAAAAGCTTTTGTAGCTCCGAGCCGTGCG 424
```



Db 5741 CCGTGTCTTCTGCTTGGCGGAGCTGCGCAAGACCTTGGCAGCTCCGAATCGTCGG 5800  
Qy 425 CCGTCGACAGCGGACGCGGACCGCCCTCTGACCAATCTCCGAGAGCGGAGAG 484  
Db 5801 CCGTCGACAGCGGACGCGGACCGCCCTCTGACCAAGCTCCGAGAGCGGAGAGCGG 5860  
Qy 485 GATCTGACGTTGATGCTATCTCTCAATGCCCCCTTGAAGGGGAGAGCGGGGAGCCCG 544  
Db 5861 GATCCGACGTTGATGCTATCTCTCAATGCCCCCTTGAAGGGGAGAGCGGGGAGTCGG 5920  
Qy 545 ATCTGACGACGCGGCTTGGTCTTACCGGTGATGAGGAGCGGGTGAAGAGTGTCTGCT 604  
Db 5921 ATCTGACGACGCGGCTTGGTCTTACCGGTGATGAGGAGCGGGTGAAGAGTGTCTGCT 5980  
Qy 605 GCTCGATGCTCAACATGAGAGAGCGGCTCTGATCAAGCCATGCGTGGAGAGAGCA 664  
Db 5981 GCTCGATGCTCAACATGAGAGAGCGGCTCTGATCAAGCCATGCGTGGAGAGAGCA 6040  
Qy 665 AGCTGCCATCAACGCGGTGAGCACTCTTGTGCTGCTCAACCAATGTCTAGGCTA 724  
Db 6041 AGCTGCCATCAACGCGGTGAGCACTCTTGTGCTGCTCAACCAATGTCTAGGCTA 6100  
Qy 725 CCAATCCCGGACGAGGACGAGGAGAGAGGTCACCTTGAAGAGCTGCAATCC 784  
Db 6101 CCAATCTCGACGAGGACGAGGAGAGAGGTCACCTTGAAGAGCTGCAATCC 6160  
Qy 785 TGGACGATCACTACAGAGAGCTGCTCAAGAGATGAGAGGCGGAGCGTCCAGTTAAG 844  
Db 6161 TGGACGATCACTACAGAGAGCTGCTCAAGAGATGAGAGGCGGAGCGTCCAGTTAAG 6220  
Qy 845 CTGAGCTTCTATCATGAGAGAGAGCTGCAAGCTGACGCCCAATTCGGCCAAATCTA 904  
Db 6221 CTGAACTTCTATCGTGGAGAGAGCTGTAAGCTGACGCCCAATTCGGCCAAATCTA 6280  
Qy 905 AATTTGGCTATGGGGGAGAGAGAGTCCGGAACCTATCCAGAGGCGCATTAACCAATCC 964  
Db 6281 AATTTGGCTATGGGGGAGAGAGAGTCCGGAACCTATCCAGAGGCGCGTTAACCAATCC 6340  
Qy 965 GCTCCGTTGGAGAGAGCTGTTGGAAGAGCACTGAAACCAACCAATTTGACACCATCATG 1024  
Db 6341 GCTCCGTTGGAGAGAGCTGTTGGAAGAGCACTGAAACCAACCAATTTGACACCATCATG 6400  
Qy 1025 CAAGAAATGAGTTTCTGCTGCTCAACAGAGAGAGAGCGGAGCGGAGCTGCTTGA 1084  
Db 6401 CAAGAAATGAGTTTCTGCTGCTCAACAGAGAGAGAGCGGAGCGGAGCTGCTTGA 6460  
Qy 1085 TGGTGTCCAGAGCTGGGGGGTCCGTTGTCGAGAGAAATGGCCCTATAGAGTGGTCT 1144  
Db 6461 TGGTGTCCAGAGCTGGGGGGTCCGTTGTCGAGAGAAATGGCCCTATAGAGTGGTCT 6520  
Qy 1145 CCAACCTCCCTCAGGCTGTGATGGGCTCTGTACGAGATTCCAGTATTTCTCTGACAGC 1204  
Db 6521 CCAACCTCCCTCAGGCTGTGATGGGCTCTGTACGAGATTCCAGTATTTCTCTGACAGC 6580  
Qy 1205 GGGTCGAGTTCTGCTGAGAGAGCTGGAATCAAGAGAGAGAGCGGCTTGGCATATG 1264  
Db 6581 GGGTCGAGTTCTGCTGAGAGAGCTGGAATCAAGAGAGAGAGCGGCTTGGCATATG 6640  
Qy 1265 ACAACCGGCTTTTGACTCAAGAGTCACTGAGATGATCCGTTAGAGAGAGTCAATTT 1324  
Db 6641 ACAACCGGCTTTTGACTCAAGAGTCACTGAGATGATCCGTTAGAGAGAGTCAATTT 6700  
Qy 1325 ATCAATGTTGTGACTTGGCCCCGGAAGCAGACAGGCGCATAGGTCCTCAGAGAGCGC 1384  
Db 6701 ATCAATGTTGTGACTTGGCCCCGGAAGCAGACAGGCGCATAGGTCCTCAGAGAGCGC 6760  
Qy 1385 TTTATATCGGGGGTCCCTGACTAATTCAAGAGAGAGAGAGTGGGCTATCGCGGTC 1444  
Db 6761 TTTATATCGGGGGTCCCTGACTAATTCAAGAGAGAGAGAGTGGGCTATCGCGGTC 6820  
Qy 1445 GCGGAGCGGCTGCTGAGAGACTAGCTGCGGTAATCCATCATGTTACTTGAAGGCT 1504

Db 6821 GCGGAGCGGCTGCTGAGAGAGAGCTGCGGTAAATACCTCAATGTACTTGAAGCGC 6880  
Qy 1505 CTGACGCTCTGTGAGCTGCAAGCTCCAGAGCTGACAGATGCTCTGTCGGAGAGCGC 1564  
Db 6881 CTGCGGCTCTGTGAGCTGCAAGCTCCAGAGCTGACAGATGCTCTGTCGGAGAGCGC 6940  
Qy 1565 TTGTGTTATCTGTGAGAGCGGGAGAACCGAGAGAGAGCGGAGCTTACGAGTCTCA 1624  
Db 6941 TTGTGTTATCTGTGAGAGCGGGAGAACCGAGAGAGAGCGGAGCTTACGAGTCTCA 7000  
Qy 1625 CGAGGCTATGACTAGTACTTGCCTCCCGGAGAGCGGCGCCCAACGAGATAGAC 1684  
Db 7001 CGAGGCTATGACTAGTACTTGCCTCCCGGAGAGCGGCGCCCAACGAGATAGAC 7060  
Qy 1685 TGGAGTTGATTAATATATGCTCTCTCAATGTGTGCTGCGGAGAGATGATCTGCGA 1744  
Db 7061 TGGAGTTGATTAATATATGCTCTCTCAATGTGTGCTGCGGAGAGATGATCTGCGA 7120  
Qy 1745 GGGTATTAATACCTCAACCGGTGACCC 1769  
Db 7121 GGGTATTAATACCTCAACCGGTGACCC 7145

Search completed: February 19, 2005, 01:26:41  
Job time : 995.878 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using SW model

Run on: February 19, 2005, 00:21:46 ; Search time 304.59 Seconds  
(without alignments)  
9508.540 Million cell updates/sec

Title: US-09-664-363-3

Sequence: 1 CAATGACTTCCAGACG.....ACTACTCACCCTACCGC 1770

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

1: Issued Patents NA: \*  
2: /cgn2\_6/prodata/1/ina/5A\_COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/5B\_COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/5B\_COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/5B\_COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1770	100.0	1770	US-08-191-160-3	Sequence 3, Appl
2	1763.6	99.6	3750	US-08-191-160-22	Sequence 22, Appl
3	1553.8	87.8	8638	US-10-029-907-7	Sequence 7, Appl
4	1553.8	87.8	8638	US-10-029-907-25	Sequence 25, Appl
5	1553.4	87.8	8642	US-10-029-907-2	Sequence 2, Appl
6	1552.2	87.7	7989	US-09-539-601-10	Sequence 10, Appl
7	1552.2	87.7	8001	US-09-539-601-4	Sequence 4, Appl
8	1552.2	87.7	8637	US-09-539-601-4	Sequence 4, Appl
9	1552.2	87.7	8638	US-10-029-907-6	Sequence 6, Appl
10	1552.2	87.7	8638	US-10-029-907-24	Sequence 24, Appl
11	1552.2	87.7	8639	US-10-029-907-1	Sequence 1, Appl
12	1552.2	87.7	8643	US-10-029-907-4	Sequence 13, Appl
13	1552.2	87.7	8649	US-09-539-601-13	Sequence 1, Appl
14	1552.2	87.7	11076	US-09-539-601-1	Sequence 1, Appl
15	1550.6	87.6	8648	US-10-029-907-5	Sequence 5, Appl
16	1549	87.5	8001	US-09-539-601-22	Sequence 22, Appl
17	1549	87.5	11076	US-09-539-601-25	Sequence 25, Appl
18	1547.4	87.4	8001	US-09-539-601-16	Sequence 16, Appl
19	1547.4	87.4	11076	US-09-539-601-19	Sequence 19, Appl
20	1545.8	87.3	8001	US-09-539-601-28	Sequence 28, Appl
21	1545.8	87.3	11076	US-09-539-601-31	Sequence 31, Appl
22	1532.2	86.6	2991	US-08-324-977-49	Sequence 49, Appl
23	1532.2	86.6	2991	US-08-324-977-49	Sequence 49, Appl
24	1532.2	86.6	2991	US-08-324-977-49	Sequence 49, Appl
25	1532.2	86.6	2991	US-08-324-977-49	Sequence 49, Appl
26	1532.2	86.6	7863	US-08-324-977-35	Sequence 35, Appl
27	1532.2	86.6	7863	US-08-324-977-35	Sequence 35, Appl

28	1532.2	86.6	7863	US-08-904-686A-35	Sequence 35, Appl
29	1532.2	86.6	7863	US-09-315-850-35	Sequence 35, Appl
30	1532.2	86.6	7917	US-08-324-977-31	Sequence 31, Appl
31	1532.2	86.6	7917	US-08-384-616-31	Sequence 31, Appl
32	1532.2	86.6	7917	US-08-904-686A-31	Sequence 31, Appl
33	1532.2	86.6	7917	US-09-315-850-31	Sequence 31, Appl
34	1532.2	86.6	9030	US-08-324-977-13	Sequence 13, Appl
35	1532.2	86.6	9030	US-08-324-977-13	Sequence 13, Appl
36	1532.2	86.6	9030	US-08-904-686A-13	Sequence 13, Appl
37	1532.2	86.6	9030	US-09-315-850-13	Sequence 13, Appl
38	1532.2	86.6	9416	US-08-324-977-1	Sequence 1, Appl
39	1532.2	86.6	9416	US-08-384-616-1	Sequence 1, Appl
40	1532.2	86.6	9416	US-08-904-686A-1	Sequence 1, Appl
41	1532.2	86.6	9416	US-09-315-850-1	Sequence 1, Appl
42	1532.2	86.6	9416	US-08-823-895A-27	Sequence 27, Appl
43	1529	86.4	9595	US-09-014-416-4	Sequence 4, Appl
44	1522.6	86.0	9472	US-08-150-204E-96	Sequence 96, Appl
45	1508.2	85.2	9413	US-09-827-688-6	Sequence 6, Appl

## ALIGNMENTS

RESULT 1  
US-08-191-160-3  
Sequence 3, Application US/08191160  
Patent No. 6210675  
GENERAL INFORMATION:  
APPLICANT: Highfield, Peter Edmund  
APPLICANT: Rodgers, Brian Colin  
APPLICANT: Tedder, Richard Seton  
APPLICANT: Barbara, John Anthony James  
TITLE OF INVENTION: Viral Agent  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, Figg, Ernst & Kurz  
STREET: 1700 K Street  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage  
COMPUTER: IBM AT compatible  
OPERATING SYSTEM: MS-DOS V3.2  
SOFTWARE: Wordperfect 5.0 (DOS text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/191,160  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/628,516  
FILING DATE: 17 DEC 1990  
APPLICATION NUMBER: UK 89 28 562.1  
FILING DATE: 18 DEC 1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 90 04 414.0  
FILING DATE: 27 FEB 1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 90 04 814.1  
FILING DATE: 03 MAR 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: E. Anthony Figg  
REGISTRATION NUMBER: 27,195  
REFERENCE/DOCKET NUMBER: 1645-103A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 833-5740  
TELEFAX: (202) 833-5740  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1770 base pairs  
TYPE: nucleotide with corresponding protein  
STRANDEDNESS: single

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; ORIGINAL SOURCE:
; ORGANISM: human; serum infectious for PT-NANBH
; IMMEDIATE SOURCE:
; LIBRARY: clone Jg2 from cDNA library in lambda gtl1
; FEATURE:
; LOCATION: from 1 to 1770 bp portion of the PT-NANBH
; LOCATION: polyprotein
; OTHER INFORMATION: probably encodes viral non-structural
; OTHER INFORMATION: proteins
us-08-191-160-3

```

```

Query Match      100.0%; Score 1770; DB 3; Length 1770;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CAAATGACTTCCAGACGCTGATCATGAGGCCAACCCTCTGTGGCGGATGATG 60
DB 1 CAAATGACTTCCAGACGCTGATCATGAGGCCAACCCTCTGTGGCGGATGATG 60
QY 61 GCGGGGACATTAACCGCGTGGAGTGCAGAAACAAGTAGTAATCCTGAGCTTTGAC 120
DB 61 GCGGGGACATTAACCGCGTGGAGTGCAGAAACAAGTAGTAATCCTGAGCTTTGAC 120
QY 121 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTTGGGAAA 180
DB 121 CCGCTCCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCCGCGGAGATCTTGGGAAA 180
QY 181 TCCAGAAATTTCCCAACAGCGATGCCCGCATGGGCAACCGCGGATTAACAACCTCCGCTG 240
DB 181 TCCAGAAATTTCCCAACAGCGATGCCCGCATGGGCAACCGCGGATTAACAACCTCCGCTG 240
QY 241 CTGAGTCCCTGGAAGGCGCCGGAACGTAACGTCCTCCAGTGTATCATGGGTGCCCATGCGCA 300
DB 241 CTGAGTCCCTGGAAGGCGCCGGAACGTAACGTCCTCCAGTGTATCATGGGTGCCCATGCGCA 300
QY 301 CCTACTAAGACCCCTCTATATACACCTCCACGGAAGAAAGAGACAGTTGTTCTGACAGA 360
DB 301 CCTACTAAGACCCCTCTATATACACCTCCACGGAAGAAAGAGACAGTTGTTCTGACAGA 360
QY 361 TCCACCGGTCTTCTTGGCCCTGCGGAGCTTGGCCAAAGGCTTTTGGTAGTCTCCGACCG 420
DB 361 TCCACCGGTCTTCTTGGCCCTGCGGAGCTTGGCCAAAGGCTTTTGGTAGTCTCCGACCG 420
QY 421 TCGGCGGTGACAGCGGCAACCGGCCCTCCGCAATTCCTCCGACGAGCGGCGGA 480
DB 421 TCGGCGGTGACAGCGGCAACCGGCCCTCCGCAATTCCTCCGACGAGCGGCGGA 480
QY 481 GGAGATCTGACGTTGAGTCTGATTCCTTCATGCCCCCTTGAAGGGGAGCCGGGGAGC 540
DB 481 GGAGATCTGACGTTGAGTCTGATTCCTTCATGCCCCCTTGAAGGGGAGCCGGGGAGC 540
QY 541 CCCGATCTGACGAGCGGGTCTTGGTCTACCGTGAAGAGAGCGGAGGAGTCTGTC 600
DB 541 CCCGATCTGACGAGCGGGTCTTGGTCTACCGTGAAGAGAGCGGAGGAGTCTGTC 600
QY 601 TGCCTCTGATGTCTTCAACATGAGACAGGCGCTGATATCAGCCATGCGTGGAGGAA 660
DB 601 TGCCTCTGATGTCTTCAACATGAGACAGGCGCTGATATCAGCCATGCGTGGAGGAA 660
QY 661 AGCAAGTGCCTCATCAAGCGCTTGAAGCACTCTTGTGCGTCAACAACATGCTTAC 720
DB 661 AGCAAGTGCCTCATCAAGCGCTTGAAGCACTCTTGTGCGTCAACAACATGCTTAC 720
QY 721 GCTACCAATCCCGAGCGCAAGCCAGCGGCAAGAAAGTCACTTTGACAGACTGCA 780
DB 721 GCTACCAATCCCGAGCGCAAGCCAGCGGCAAGAAAGTCACTTTGACAGACTGCA 780
QY 781 ATCCGAGAGATCATCAAGAGACGTCAGAGAGATGAAGGCAAGCGCTCCACAGTT 840
DB 781 ATCCGAGAGATCATCAAGAGACGTCAGAGAGATGAAGGCAAGCGCTCCACAGTT 840

```

```

QY 841 AAGGCTAAGCTTCTATCATGAGAGAAAGCAAGCTGCAAGCTGAGCGCCCACTTGGCCAAA 900
DB 841 AAGGCTAAGCTTCTATCATGAGAGAAAGCAAGCTGCAAGCTGAGCGCCCACTTGGCCAAA 900
QY 901 TCTAAATTTGGCTATGAGGAGCAAGAGCTCCGGAACCTATCAGCAAGGCCATTAACAC 960
DB 901 TCTAAATTTGGCTATGAGGAGCAAGAGCTCCGGAACCTATCAGCAAGGCCATTAACAC 960
QY 961 ATCCGCTCCGTGTGGAGAGACTTGTGGAAGACACTGAAACACCAATTGACACCACTC 1020
DB 961 ATCCGCTCCGTGTGGAGAGACTTGTGGAAGACACTGAAACACCAATTGACACCACTC 1020
QY 1021 ATGCAAAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGAGGCGCGCAACGCTTGC 1080
DB 1021 ATGCAAAAAATGAGGTTTCTGCGTCCAAACAGAGAGAGAGGCGCGCAACGCTTGC 1080
QY 1081 CTATTCGTGTTCCCAAGACTTGGGGGTCCGTGTGCGGAAATGGCCCTTATGACGTG 1140
DB 1081 CTATTCGTGTTCCCAAGACTTGGGGGTCCGTGTGCGGAAATGGCCCTTATGACGTG 1140
QY 1141 GTCTCAACCTCCCTCAGAGCTGTGATGGGCTCCTCGTACGGAATTCAGTATTTCTTGA 1200
DB 1141 GTCTCAACCTCCCTCAGAGCTGTGATGGGCTCCTCGTACGGAATTCAGTATTTCTTGA 1200
QY 1201 CAGCGGAGTGAATTCCTGAGTGAACGCTGGAATCAAGAAACCCCTATGAGGCTTTGCA 1260
DB 1201 CAGCGGAGTGAATTCCTGAGTGAACGCTGGAATCAAGAAACCCCTATGAGGCTTTGCA 1260
QY 1261 TATGACACCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 1320
DB 1261 TATGACACCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 1320
QY 1321 ATTATCAATGTTGACTTGAGCTTGAGCCCGGCAAGCAGGCGCATAGGTCGCTCAGAG 1380
DB 1321 ATTATCAATGTTGACTTGAGCTTGAGCCCGGCAAGCAGGCGCATAGGTCGCTCAGAG 1380
QY 1381 CGGCTTATATCGGGGAGTCCCTGATTAATCAAAAGGGGCAAGCTGGCGCTATCGCGG 1440
DB 1381 CGGCTTATATCGGGGAGTCCCTGATTAATCAAAAGGGGCAAGCTGGCGCTATCGCGG 1440
QY 1441 TCCCGCGGAGCGGCTGTGACGACTAGCTAGCTGCGGTATACCTTCACTGTTACTTGAAG 1500
DB 1441 TCCCGCGGAGCGGCTGTGACGACTAGCTAGCTGCGGTATACCTTCACTGTTACTTGAAG 1500
QY 1501 GCTCTGACGCTGTGACGCTGCAAGAGCTCCAGGACTGCAAGTCTGTGTGCGAGAGC 1560
DB 1501 GCTCTGACGCTGTGACGCTGCAAGAGCTCCAGGACTGCAAGTCTGTGTGCGAGAGC 1560
QY 1561 GGCCTTGTCTTATCTGTGAGAGCGCGGAAACCAAGAGAGCGCGCGAGCTTACGAGTC 1620
DB 1561 GGCCTTGTCTTATCTGTGAGAGCGCGGAAACCAAGAGAGCGCGCGAGCTTACGAGTC 1620
QY 1621 TTCAAGAGGCTATGACTAGTACTTGTGCCCCCCCCCGGGGACCGGCCCAACCAAGATAC 1680
DB 1621 TTCAAGAGGCTATGACTAGTACTTGTGCCCCCCCCCGGGGACCGGCCCAACCAAGATAC 1680
QY 1681 GACCTGAGTGTATTAACATCATGCTCCCAATGTGTGCGGCGGAGATGATCTGGC 1740
DB 1681 GACCTGAGTGTATTAACATCATGCTCCCAATGTGTGCGGCGGAGATGATCTGGC 1740
QY 1741 AAAAGGGTATTAATCACTCAACCCGAGACCG 1770
DB 1741 AAAAGGGTATTAATCACTCAACCCGAGACCG 1770

```

```

RESULT 2
us-08-191-160-22
; Sequence 22, Application US/08191160
; Patent No. 6210675
; GENERAL INFORMATION:
; APPLICANT: Highfield, Peter Edmund
; APPLICANT: Rodgers, Brian Colin
; APPLICANT: Tedder, Richard Seton

```

APPLICANT: Barbara, John Anthony James  
 TITLE OF INVENTION: Viral Agent  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Rothwell, Figg, Ernst & Kurz  
 STREET: 1700 K Street  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage  
 COMPUTER: IBM AT compatible  
 OPERATING SYSTEM: MS-DOS V3.2  
 SOFTWARE: Wordperfect 5.0 (DOS text)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/191,160  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/628,516  
 FILING DATE: 17 DEC 1990  
 APPLICATION NUMBER: UK 89 28 562.1  
 FILING DATE: 18 DEC 1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: UK 90 04 414.0  
 FILING DATE: 27 FEB 1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: UK 90 04 814.1  
 FILING DATE: 03 MAR 1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: E. Anthony Figg  
 REGISTRATION NUMBER: 27,195  
 REFERENCE/DOCKET NUMBER: 1645-103A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 833-5740  
 TELEFAX: (202) 833-5744  
 INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3750 base pairs  
 TYPE: nucleotide with corresponding protein  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to genomic RNA  
 ORIGINAL SOURCE:  
 ORGANISM: human; serum infectious for PT-NANBH  
 IMMEDIATE SOURCE:  
 LIBRARY: cDNA clones from 3' end of the genome  
 FEATURE:  
 LOCATION: from 1 to 3750 bp portion of the PT-NANBH  
 LOCATION: polyprotein  
 OTHER INFORMATION: viral non-structural proteins  
 US-08-191-160-22

Query Match 99.8%; Score 1763.6; DB 3; Length 3750;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1766; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACTCTGTGCGGAGATGAGATG 60  
 DB 1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACTCTGTGCGGAGATGAGATG 2040  
 QY 61 GGGGGGAGATTACCCGCGTGAAGTCAAGAACAGGTAACTCTGACCTTTTGAC 120  
 DB 2041 GGGGGGAGATTACCCGCGTGAAGTCAAGAACAGGTAACTCTGACCTTTTGAC 2100  
 QY 121 CCGCTCCGAGCGAGAGATGAGCGGAGATGTCCTCCGCGGAGATCCGCGGAAA 180  
 DB 2101 CCGCTCCGAGCGAGAGATGAGCGGAGATGTCCTCCGCGGAGATCCGCGGAAA 2160  
 QY 181 TCCAGAAATTTCCACAGCGATGCCGAGATGGGCAAGCCGAGATTCAACCTCCGCTG 240  
 DB 2161 TCCAGAAATTTCCACAGCGATGCCGAGATGGGCAAGCCGAGATTCAACCTCCGCTG 2220

QY 241 CTGAGATCTTGAAAGGCCCCGAGACTACGTCCTCCAGTGTATATGGGTGCCCCACTGCCA 300  
 DB 2221 CTGAGATCTTGAAAGGCCCCGAGACTACGTCCTCCAGTGTATATGGGTGCCCCACTGCCA 2280  
 QY 301 CCTACTAAGACCCCTCTCTATACACCTCCACGAGAGAAAGAGACATGTTCTTGACAAA 360  
 DB 2281 CCTACTAAGACCCCTCTCTATACACCTCCACGAGAGAAAGAGACATGTTCTTGACAAA 2340  
 QY 361 TCCACCGTGTCTTTCGCTCCGAGAGCTTGCCCAAAAGCTTTTGTAGCTTCGGAACG 420  
 DB 2341 TCCACCGTGTCTTTCGCTCCGAGAGCTTGCCCAAAAGCTTTTGTAGCTTCGGAACG 2400  
 QY 421 TGGGCGGTGAGACAGGGGCAAGGCAACCGCCCTCCGACCAATTCCTCCGAGAGCGGGA 480  
 DB 2401 TGGGCGGTGAGACAGGGGCAAGGCAACCGCCCTCCGACCAATTCCTCCGAGAGCGGGA 2460  
 QY 481 GCAGATCTGAGGTTGAGTGTATTCTTCATGCCCCCTTTGAGGGGAGCCGGGGAGC 540  
 DB 2461 GCAGATCTGAGGTTGAGTGTATTCTTCATGCCCCCTTTGAGGGGAGCCGGGGAGC 2520  
 QY 541 CCGATCTCAGGAGCGGCTTGTGTCTACCGTGAAGTGAAGGCGGCTGAGACCTGTC 600  
 DB 2521 CCGATCTCAGGAGCGGCTTGTGTCTACCGTGAAGTGAAGGCGGCTGAGACCTGTC 2580  
 QY 601 TGTGCTGATGTCTTCAACATGAGACAGGCGCTGTATACGCGCANTGGCGTGGAGAAA 660  
 DB 2581 TGTGCTGATGTCTTCAACATGAGACAGGCGCTGTATACGCGCANTGGCGTGGAGAAA 2640  
 QY 661 AGCAAGCTCCCATCAACGCGTGTGAGCACTTTGCTCGTACACCAACATGCTCTAC 720  
 DB 2641 AGCAAGCTCCCATCAACGCGTGTGAGCACTTTGCTCGTACACCAACATGCTCTAC 2700  
 QY 721 GCTACCAATCCCGCAGCGCAAGCCAGCGGCAAGAGATGACCTTTGACAGACTGCAA 780  
 DB 2701 GCTACCAATCCCGCAGCGCAAGCCAGCGGCAAGAGATGACCTTTGACAGACTGCAA 2760  
 QY 781 ATCCGTGAGATCACTACAGAGACGCTCAAGAGATGAAGGCGGACGCTCCAGATT 840  
 DB 2761 ATCCGTGAGATCACTACAGAGACGCTCAAGAGATGAAGGCGGACGCTCCAGATT 2820  
 QY 841 AAGCTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACGCCCCACATTCGGCCAAA 900  
 DB 2821 AAGCTAAGCTTCTATCACTAGAGAGAGCTGCAAGCTGACGCCCCACATTCGGCCAAA 2880  
 QY 901 TCTAAATTTGGCTATGAGGCAAGAGAGCTCCGGAACCTATCCAGCAAGGCATTAAACAC 960  
 DB 2881 TCTAAATTTGGCTATGAGGCAAGAGAGCTCCGGAACCTATCCAGCAAGGCATTAAACAC 2940  
 QY 961 ATCCGCTCCGTGTGGAGAGACTTGTGGAAGACACTGAAACACCAATTGACACACATC 1020  
 DB 2941 ATCCGCTCCGTGTGGAGAGACTTGTGGAAGACACTGAAACACCAATTGACACACATC 3000  
 QY 1021 ATGGCAAAATATGAGGTTTTCGCTGCAACCAAGAGAGAGAGCGCAAGCCAGCTCCG 1080  
 DB 3001 ATGGCAAAATATGAGGTTTTCGCTGCAACCAAGAGAGAGAGCGCAAGCCAGCTCCG 3060  
 QY 1081 CTATGCTGTCCAGACTTGGGGGTCGCTGTGTGCGAAGAAATGAGCCCTCTATGACGTC 1140  
 DB 3061 CTATGCTGTCCAGACTTGGGGGTCGCTGTGTGCGAAGAAATGAGCCCTCTATGACGTC 3120  
 QY 1141 GTTTCACCTCCCTTACAGCTGTGATGAGGCTCTCTGATAGGATTCAGATTCTCTGGA 1200  
 DB 3121 GTTTCACCTCCCTTACAGCTGTGATGAGGCTCTCTGATAGGATTCAGATTCTCTGGA 3180  
 QY 1201 CAGCGGCTGAGATTCTCTGTTGAAGCTGGAATCAAGAGAGAGCCCTATGAGCTTTGGA 1260  
 DB 3181 CAGCGGCTGAGATTCTCTGTTGAAGCTGGAATCAAGAGAGAGCCCTATGAGCTTTGGA 3240  
 QY 1261 TATGACACCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 1320  
 DB 3241 TATGACACCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCA 3300

QY 1321 ATTATCAATGTTGTGACTTGGCCCCCGAAGCCAGACAGTCAAGTCTGCTACAGAG 1380  
|  
|  
|  
Db 3301 ATTATCAATGTTGTGACTTGGCCCCCGAAGCCAGACAGTCAAGTCTGCTACAGAG 3360  
|  
|  
|  
QY 1381 CGGCTTTAATGCGGGGTCCCTGACTAATTCAAAAAGGCGAGAACTGCGGCTATGCGCG 1440  
|  
|  
|  
Db 3361 CGGCTTTAATGCGGGGTCCCTGACTAATTCAAAAAGGCGAGAACTGCGGCTATGCGCG 3420  
|  
|  
|  
QY 1441 TCCCGCGAGCGCGCTGCTGACGACTAGCTCGGTAATACCTCAATGTTACTTGAAG 1500  
|  
|  
|  
Db 3421 TCCCGCGAGCGCGCTGCTGACGACTAGCTCGGTAATACCTCAATGTTACTTGAAG 3480  
|  
|  
|  
QY 1501 GCCTTGACGCTGCTGACCTGCAAGCTCCAGAGCTGACGATGCTGTGTCGAGAGC 1560  
|  
|  
|  
Db 3481 GCCTTGACGCTGCTGACCTGCAAGCTCCAGAGCTGACGATGCTGTGTCGAGAGC 3540  
|  
|  
|  
QY 1561 GGCCTTGCTGATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACGAGTC 1620  
|  
|  
|  
Db 3541 GGCCTTGCTGATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACGAGTC 3600  
|  
|  
|  
QY 1621 TTCAGGAGGCTATGACTAGTACTCTGCCCCCGGAGACCCGCCCCAAGCAATAC 1680  
|  
|  
|  
Db 3601 TTCAGGAGGCTATGACTAGTACTCTGCCCCCGGAGACCCGCCCCAAGCAATAC 3660  
|  
|  
|  
QY 1681 GACTGAGAGTATTAACATCATGCTCTCCAAATGTTGCTGCGCAGCAGATCTGAGC 1740  
|  
|  
|  
Db 3661 GACTGAGAGTATTAACATCATGCTCTCCAAATGTTGCTGCGCAGCAGATCTGAGC 3720  
|  
|  
|  
QY 1741 AAAAGGTTACTACTCTCAACCCGTGACCG 1770  
|  
|  
|  
Db 3721 AAAAGGTTACTACTCTCAACCCGTGACCG 3750  
|  
|  
|

## RESULT 3

US-10-029-907-7  
; Sequence 7, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PabctSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802)... (8407)  
US-10-029-907-7

Query Match 87.8%; Score 1553.8; DB 4; Length 8638;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1633; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 5 ATGACTTCCAGACGCTGACCTCATCGAGCCAACTCTCTGCGGCGATGAGATGAGCG 64  
|  
|  
|  
Db 6030 ATGACTCCCCGAGCGCTGACCTCATCGAGCCAACTCTCTGCGGCGAGATGAGCG 6089  
|  
|  
|  
QY 65 GGGACATTACCCCGTGGAGTCAAGAGACAAAGTATGTAATCTTGAATCTTTGACCGCG 124  
|  
|  
|  
Db 6090 GGAATCATCACCCCGTGGAGTCAAGAGAAATAGTATGTAATTTTGAATCTTTGACCGCG 6149  
|  
|  
|  
QY 125 TCCGAGCGGAGAGATGAGCGGGAAGTCCGTCGCCGCGAGATCCGCGGAAATCCA 184  
|  
|  
|  
Db 6150 TCCAGCGGAGAGATGAGCGGGAAGTCCGTCGCCGCGAGATCCGCGGAGTCCA 6209  
|  
|  
|

QY 185 AGAAATTCACACAGCGATGCGCGATGAGGACCGCCGGAATTACAACTCCGCTGCTGG 244  
|  
|  
|  
Db 6210 GGAATTCCTCGAGCGATGCGCGATGAGGACCGCCGGAATTACAACTCCGCTGCTGAG 6269  
|  
|  
|  
QY 245 AGTCTGGAAGAGCCCGGACTATAGTCTCTCCAGTGTATCATGGGTGCCACTGCCACTA 304  
|  
|  
|  
Db 6270 AGTCTGGAAGAGCCCGGACTATAGTCTCTCCAGTGTATCATGGGTGCCACTGCCACTG 6329  
|  
|  
|  
QY 305 CTAAAGACCCCTCTTATACACCTCCACGGAAGAGAGGACAGTTGTTGACAGAAATCCA 364  
|  
|  
|  
Db 6330 CCAAGGCCCTCCGATACCACTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6389  
|  
|  
|  
QY 365 CGGTCTCTTGCCTCGGAGAGCTTGCACAAAGGCTTTTGTAGCTCCGAGCCGTGCG 424  
|  
|  
|  
Db 6390 CGGTCTCTTGCCTCGGAGAGCTTGCACAAAGGCTTGCAGAGCTCCGAAATGTGCG 6449  
|  
|  
|  
QY 425 CCGTGAAGAGGAG 484  
|  
|  
|  
Db 6450 CCGTGAAGAGGAG 6509  
|  
|  
|  
QY 485 GATCTGAGCTTGAAGTCTGATTCCTCCATGCCCCCTTGAAGGAGAGAGAGAGAGAGAG 544  
|  
|  
|  
Db 6510 GATCTGAGCTTGAAGTCTGATTCCTCCATGCCCCCTTGAAGGAGAGAGAGAGAGAG 6569  
|  
|  
|  
QY 545 ATCTGAGGAGAGGCTTGTGCTTACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604  
|  
|  
|  
Db 6570 ATCTGAGGAGAGGCTTGTGCTTACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6629  
|  
|  
|  
QY 605 GCTGATGCTCTTACATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664  
|  
|  
|  
Db 6630 GCTGATGCTCTTACATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6689  
|  
|  
|  
QY 665 AGCTGCCCATCAAGCGGCTTGAAGCAACTCTTGTGCTGCTGACCAACAATGTTACGCTTA 724  
|  
|  
|  
Db 6690 AGCTGCCCATCAAGCGGCTTGAAGCAACTCTTGTGCTGCTGACCAACAATGTTACGCTTA 724  
|  
|  
|  
QY 725 CCACATCCCGAGAGCGCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784  
|  
|  
|  
Db 6750 CCACATCCCGAGAGCGGCTTGAAGCAACTCTTGTGCTGCTGACCAACAATGTTACGCTTA 784  
|  
|  
|  
QY 785 TGAAGCATCATTACAG 844  
|  
|  
|  
Db 6810 TGAAGCATCATTACAG 844  
|  
|  
|  
QY 845 CTAAGCTTCTATCGTGAAG 904  
|  
|  
|  
Db 6870 CTAAGCTTCTATCGTGAAG 904  
|  
|  
|  
QY 905 AATTGGCTATGAGGCAAG 964  
|  
|  
|  
Db 6930 AATTGGCTATGAGGCAAG 964  
|  
|  
|  
QY 965 GCTCCGTGTGAG 1024  
|  
|  
|  
Db 6990 GCTCCGTGTGAG 1024  
|  
|  
|  
QY 1025 CAAAAAATGAG 1084  
|  
|  
|  
Db 7050 CAAAAAATGAG 1084  
|  
|  
|  
QY 1085 TCGTGTCCAGAGCTTGGGAG 1144  
|  
|  
|  
Db 7110 TCGTGTCCAGAGCTTGGGAG 1144  
|  
|  
|  
QY 1145 CCACCTTCCCTCAG 1204  
|  
|  
|  
Db 7170 CCACCTTCCCTCAG 1204  
|  
|  
|  
QY 1205 GGGTGAAGTCTCTGAG 1264  
|  
|  
|  
Db 7230 GGGTGAAGTCTCTGAG 1264  
|  
|  
|  
QY 1265 ACACCCGCTGTTTGAAGTCAAGTCACTGAAGATGACATCCGTGAGAGAGTCAATTT 1324  
|  
|  
|

Db 7290 AGACCCGCTGTTTGACTCAAGCGTCACTGAGAAATACATCCCTGTGAGAGTCAATCT 7349  
Qy 1325 ATCAATGTTGTGACTTGGCCCCCGAAGCAAGACGAGCCATAGTCTCTCAAGAGCGG 1384  
Db 7350 ACCAATGTTGTGACTTGGCCCCCGAAGCAAGACGAGCCATAGTCTCTCAAGAGCGG 7409  
Qy 1385 TTATATCGGGGGTCCCTGAGCTAATTTCAAAAGGAGAACTGGCGCTATCGCGGCTGCC 1444  
Db 7410 TTTACATCGGGGGCCCCCTGACTAATTTCAAAAGGAGAACTGGCGCTATCGCGGCTGCC 7469  
Qy 1445 GCGGAGCGGCGCTGCTGAGCACTAGTGGGTATACCTTCACATGTTACTTGAAGGCT 1504  
Db 7470 GCGGAGCGGCTGCTGAGCACTAGTGGGTATACCTTCACATGTTACTTGAAGGCT 7529  
Qy 1505 CTGACGCTGTGAGCTGCAAGAGCTCAAGAGCTGCAAGTGTGTGTGCGAGAGCGGCT 1564  
Db 7530 CTGCGGCTGTGAGCTGCAAGAGCTCAAGAGCTGCAAGTGTGTGTGCGAGAGCGGCT 7589  
Qy 1565 TTGTGTTATCTGTGAGAGCGGGAACCAAGAGAGCGGGAAGCTTACGATCTTCA 1624  
Db 7590 TTGTGTTATCTGTGAGAGCGGGAACCAAGAGAGCGGGAAGCTTACGATCTTCA 7649  
Qy 1625 CCGAGGCTATGACTAGTACTGTGCCCCCGGAGACCGGCCCAACGAGATTAAGACT 1684  
Db 7650 CCGAGGCTATGACTAGTACTGTGCCCCCGGAGACCGGCCCAACGAGATTAAGACT 7709  
Qy 1685 TGGAGTTGATACATCATGCTCTCCATGTGTGCGTGGGAGCATGATCTGGCAAAA 1744  
Db 7710 TGGAGTTGATACATCATGCTCTCCATGTGTGCGTGGGAGCATGATCTGGCAAAA 7769  
Qy 1745 GGGTATACTACTCCTACCCGCTGACCC 1769  
Db 7770 GGGTATACTACTCCTACCCGCTGACCC 7794

## RESULT 4

US-10-029-907-25  
Sequence 25, Application US/10029907  
Patent No. 6706874  
GENERAL INFORMATION:  
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
TITLE OF INVENTION: HEPATITIS C VIRUS  
FILE REFERENCE: 13/083  
CURRENT APPLICATION NUMBER: US/10/029,907  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/257,857  
PRIOR FILING DATE: 2000-12-22  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 25  
LENGTH: 8638  
TYPE: DNA  
ORGANISM: HCV  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1802) ... (8407)  
US-10-029-907-25

Query Match 87.8%; Score 1553.8; DB 4; Length 8638;  
Best Local Similarity 92.5%; Pired. No. 0;  
Matches 1633; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
Qy 5 ATGACTTCCAGACGCTGACCTCATGAGGCCAACCCTCTGTGGCGGACATGAGATGGCG 64  
Db 6030 ATGACTTCCAGACGCTGACCTCATGAGGCCAACCCTCTGTGGCGGAGATGGCG 6089  
Qy 65 GGGACATTAACCCGCTGAGTCAAGAAACAAGTATGATCTTTCAGACCCG 124  
Db 6090 GGAACATCAACCCGCTGAGTCAAGAAACAAGTATGATCTTTCAGACCCG 6149  
Qy 125 TCCGACGAGAGAGATGAGCGGGAAGTGTCCCTCCCGGAGAGATCTTCGGAATCCA 184

Db 6150 TCCAGCGGAGAGATGAGAGGAAGTATCCGTTCCGCGGAGATCTCCGAGGTTCA 6209  
Qy 185 AGAATTTCCACGAGGATGCCCGATGGGACCGCCGGATTACAACTCCGCTCTGG 244  
Db 6210 GGAATTTCCGCTGAGGATGCCCGATGGGACCGCCGGATTACAACTCCGCTCTGG 6269  
Qy 245 AGTCCGGAAGGCCCCCGGACTAGTCCCTCCAGTGGTATAGGGTCCGACGCACTA 304  
Db 6270 AGTCCGGAAGGCCCCCGGACTAGTCCCTCCAGTGGTATAGGGTCCGACGCTTGG 6329  
Qy 305 CTAAGACCCCTCTATACCACTTCCACGAGAAAGAGACAGTGTGTTGAGAGATCCA 364  
Db 6330 CCAAGGCCCCCTCCGATACCACTTCCACGAGAAAGAGACAGTGTGTTGAGAGATCCA 6389  
Qy 365 CCGTGTCTTCTGCTGCGGAGCTTGGCAAAAGCTTTGTGTAGCTCCGACGCTCG 424  
Db 6390 CCGTGTCTTCTGCTGCGGAGCTTGGCAAAAGCTTTGTGTAGCTCCGACGCTCG 6449  
Qy 425 CCGTGCAGCGGCAAGGCAACCGCCCTCTGACCAATCCCTCCGAGCGAGGAGCAG 484  
Db 6450 CCGTGCAGCGGCAAGGCAACCGCCCTCTGACCAATCCCTCCGAGCGAGGAGCAG 6509  
Qy 485 GATCTGACGTTGAGTGTATATCTCCATGCCCCCTTGAAGGAGGAGCGGAGACCCG 544  
Db 6510 GATCTGACGTTGAGTGTATATCTCCATGCCCCCTTGAAGGAGGAGCGGAGACCCG 6569  
Qy 545 ATCTCAGCAGCGGCTTGTGCTTACCGTGAAGAGGCGGTGAGAGCGTGTCTGCT 604  
Db 6570 ATCTCAGCAGCGGCTTGTGCTTACCGTGAAGAGGCGGTGAGAGCGTGTCTGCT 6629  
Qy 605 GCTTCATGTTCTTACATATGAGAGGCTCTGATACCGCATGCGGTGGGAGGAAGA 664  
Db 6630 GCTTCATGTTCTTACATATGAGAGGCTCTGATACCGCATGCGGTGGGAGGAAGA 6689  
Qy 665 AGTGGCCATCAAGCGGTTGAGCAACTCTTGTGCTGCTGACCAAACTGTCTACGCTA 724  
Db 6690 AGTGGCCATCAAGCGGTTGAGCAACTCTTGTGCTGCTGACCAAACTGTGTATGCTA 6749  
Qy 725 CCAACATCCGAGCGGCAAGCGGCAAGAGAGGTCACCTTTCAGACTGCAAAATCC 784  
Db 6750 CCAACATCCGAGCGGCAAGCGGCAAGAGAGGTCACCTTTCAGACTGCAAGTCC 6809  
Qy 785 TGGAGCATCACTTACAGAGCGTCTCAAGAGATGAAGGCGAAGGCTCCACATTAAG 844  
Db 6810 TGGAGCATCACTTACAGAGCGTCTCAAGAGATGAAGGCGAAGGCTCCACATTAAG 849  
Qy 845 CTAAGCTTCTATCAGTGAAGAGCGTCAAGCTGACGCCCCCACTTGGCCAAATCTA 904  
Db 6870 CTAAGCTTCTATCAGTGAAGAGCGTCAAGCTGACGCCCCCACTTGGCCAAATCTA 6929  
Qy 905 AATTGGCTATGAGGCAAGAGCGTCCGGAACCTATCCAGAGGCGCATTAACCATCC 964  
Db 6930 AATTGGCTATGAGGCAAGAGCGTCCGGAACCTATCCAGAGGCGCATTAACCATCC 6989  
Qy 965 GCTCCGTTGGAGAGACTTGTGGAAGACACTGAAACACCAATTGACACCACTATAG 1024  
Db 6990 GCTCCGTTGGAGAGACTTGTGGAAGACACTGAAACACCAATTGACACCACTATAG 7049  
Qy 1025 CAAAAATGAGGTTTCTGCTCCAAACGAGAGAGAGGCGCAAGCAGCTGCGCTTA 1084  
Db 7050 CAAAAATGAGGTTTCTGCTCCAAACGAGAGAGAGGCGCAAGCAGCTGCGCTTA 7109  
Qy 7110 TGTATTTCCAGATTTGGGGGTTGCTGTGCGAGAAATGAGCCCTTTCAGATGTGTCT 7169  
Db 1145 CCACTCTCCCTCAGGTTGATGAGGCTCTCTGATCGGATTTCCAGATTTCTCTGAGCAGC 1204  
Qy 7170 CCACTCTCCCTCAGGCTGATGAGGCTCTCTGATCGGATTTCCAGATTTCTCTGAGCAGC 7229  
Qy 1205 GGTGAGTTCTCTGAGAGCGCTGGAATCAAAAGAGACCCCTATGAGGCTTTCATATG 1264



Db 7230 GGGTCAGTTCCTGTGTAATGCTGGAAAGCAAGAAATGCCCTATGGGCTTCGATATAG 7289  
Qy 1265 ACACCGCGCTGTTTGTGACTCAACAGTCACTGAGATGACATCCGTGAGAGATCAATT 1324  
Db 7290 AACCCGCTGTTTGTGACTCAACAGTCACTGAGATGACATCCGTGAGAGATCAATT 7349  
Qy 1325 ATCAATGTTGTGACTTGGCCCCCGAAGCCAGACAGCCATTAAGTCTGCTACAGAGCCG 1384  
Db 7350 ACCAATGTTGTGACTTGGCCCCCGAAGCCAGACAGCCATTAAGTCTGCTACAGAGCCG 7409  
Qy 1385 TTTATATGCGGGGGTCCCTGACTTAATTCAAAAGGAGAACTCCGCTATCCCGGTGCC 1444  
Db 7410 TTATACATCGGGGGCCCCCTGACTTAATTCAAAAGGAGAACTCCGCTATCCCGGTGCC 7469  
Qy 1445 GCGCAGCGCGCTGTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1504  
Db 7470 GCGCAGCGCGTGTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 7529  
Qy 1505 CTGCAAGCTGTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1564  
Db 7530 CTGCGCGCTGTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 7589  
Qy 1565 TTGTGTTATCTGTGAGAGCGCGGGAAACCAAGAGAGCGCGCGCTTACAGATCTTCA 1624  
Db 7590 TTGTGTTATCTGTGAGAGCGCGGGAAACCAAGAGAGCGCGCGCTTACAGATCTTCA 7649  
Qy 1625 CGAGGCTATGACTAGTACTCTGCCCCCGCGGGACCCGCCCAACAGATAGAGAC 1684  
Db 7650 CGAGGCTATGACTAGTACTCTGCCCCCGCGGGACCCGCCCAACAGATAGAGAC 7709  
Qy 1685 TGGAGTTGATTAATCATCATGCTCTCCATATGTGTGCGTGGGAGAGATCATGAGCAAAA 1744  
Db 7710 TGGAGTTGATTAATCATCATGCTCTCCATATGTGTGCGTGGGAGAGATCATGAGCAAAA 7769  
Qy 1745 GGGTATACCTCACTCAACCCGTGACCC 1769  
Db 7770 GGGTATACCTCACTCAACCCGTGACCC 7794

## RESULT 5

US-10-029-907-2  
Sequence 2, Application US/10029907  
Patent No. 6706874  
GENERAL INFORMATION:  
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
FILE REFERENCE: 13/083  
CURRENT APPLICATION NUMBER: US/10/029,907  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/257,857  
PRIOR FILING DATE: 2000-12-22  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 8642  
TYPE: DNA  
ORGANISM: HCV  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1802)...(8407)  
NAME/KEY: variation  
LOCATION: 6268  
OTHER INFORMATION: r = a or g  
NAME/KEY: variation  
LOCATION: 4446  
OTHER INFORMATION: r = a or g  
US-10-029-907-2

Query Match 87.8% Score 1553.4; DB 4; Length 8642;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1632; Conservative 1; Mismatches 132; Indels 0; Gaps 0;

Qy 5 ATGACTTCCAGACGCTGACCTCATGAGCCAACTCTGTGGCGGATGAGATGGCG 64  
Db 6030 ATGACTTCCCGGAGCGCTGACCTCATGAGCCAACTCTGTGGCGGAGATGGCG 6089  
Qy 65 GGGACATTACCGCGGTGAGTCAAGAAACAAGTGTATCTGTGACTTTTCCAGCCG 124  
Db 6090 GGAACATACCGCGGTGAGTCAAGAAACAAGTGTATCTGTGACTTTTCCAGCCG 6149  
Qy 125 TCCGAGCGGAGAGATGAGCGGGAGTGTCCGTCCGCGGAGATCTCTGCGAAATCCA 184  
Db 6150 TCCAGCGGAGAGATGAGCGGGAGTGTCCGTCCGCGGAGATCTCTGCGAGTCCA 6209  
Qy 185 AGAAATTCACACAGCATGCTCCGATGGGACCGCCGATTAACAACCTCCGCTGCG 244  
Db 6210 GGAATTCCTCCGAGCATGCTCCATATGGGACCGCCGATTAACAACCTCCGCTGCG 6269  
Qy 245 AGTCTGGAAGGCCCGGACTTACGTCTCTCCAGTGTATCATGGGTGCCCATGCCACTTA 304  
Db 6270 AGTCTGGAAGGACCGGACTTACGTCTCTCCAGTGTATCATGGGTGCCCATGCCACTTG 6329  
Qy 305 CTAAAGACCCCTCTATACACCTTCAACGAGAAAGAGACAGTTGTGACAGAAATCCA 364  
Db 6330 CCAAGGCCCTCTCCATACACCTTCAACGAGAAAGAGACAGTTGTCTGACAAATCTTA 6389  
Qy 365 CCGTGTCTTTCGCTGCGGAGCTTGGCAAAAAGCTTTTGTAGCTCCGACCGTGG 424  
Db 6390 CCGTGTCTTTCGCTGCGGAGCTTGGCAAAAAGCTTCCGACAGCTCCGAATGTGCG 6449  
Qy 425 CCGTGCAGACGGGACGGCAACCGCCCTCTGACCAATCTCTGACAGAGCGGAGAG 484  
Db 6450 CCGTGCAGACGGGACGGCAACCGCCCTCTGACCAACCTCTGACAGAGCGGAGAG 6509  
Qy 485 GATCTGAGCTTGAAGTGTATCTCCATGCCCCCTTGAAGGGAGAGCGGGAGACCCG 544  
Db 6510 GATCTGAGCTTGAAGTGTATCTCCATGCCCCCTTGAAGGGAGAGCGGGAGATCCG 6569  
Qy 545 ATCTCAGGACGGGTCTTGTGTCTACCGTGAAGAGAGCGGTGAGGACGTGTCTGCT 604  
Db 6570 ATCTCAGGACGGGTCTTGTGTCTACCGTGAAGAGAGCGGTGAGGACGTGTCTGCT 6629  
Qy 605 GCTGATGTCTTACATATGACAGCGCTCTGATCAAGCTATGGCTGCGGAGAGAGCA 664  
Db 6630 GCTGATGTCTTACATATGACAGCGCGCTGATCAAGCTATGGCTGCGGAGAGAGCA 6689  
Qy 665 AGTGGCCATCAAGCGCTTGAAGCACTTGTGCGTCAACCAACATGCTTACAGCTTA 724  
Db 6690 AGTGGCCATCAAGCGCTTGAAGCACTTGTGCGTCAACCAACATGCTTACAGCTTA 6749  
Qy 725 CCAATCCCGGAGCGCAAGCGGAGAGAAAGTCACTTTGACAGACTGCAATCC 784  
Db 6750 CCAATCTCGGAGCGCAAGCGGAGAGAAAGTCACTTTGACAGACTGCAATCC 6809  
Qy 785 TGAACATCACTTACAGAGCGTGTCAAGAGATGAAGGCGAAGCGTCCCACTTAAGG 844  
Db 6810 TGAACATCACTTACAGAGCGTGTCAAGAGATGAAGGCGAAGCGTCCCACTTAAGG 6869  
Qy 845 CTAAGCTTCTATCACTGAGAGAGCGTGAAGCGCCCACTTCCGCAATCC 904  
Db 6870 CTAAGCTTCTATCGTGAAGAGCGTGAAGCGCCCACTTCCGCAATCC 6929  
Qy 905 AATTGGCTATGGGCAAGAGCGTCCGAACCTATCCAGAGAGCCATTAACCAATCC 964  
Db 6930 AATTGGCTATGGGCAAGAGCGTCCGAACCTATCCAGAGAGCCATTAACCAATCC 6989  
Qy 965 GCTCCGTGTGGAGAGCTTGTGAAGACACTGAACCAATTTGACCACTCATGG 1024  
Db 6990 GCTCCGTGTGAAGAGCTTGTGAAGACACTGAAGCACTGAAGCACTCATCATGG 7049  
Qy 1025 CAAAAAATGAGGTTTTCGCTCCACCAAGAGAGAGCGGAGCGCACTGCGCTTA 1084  
Db 7050 CAAAAAATGAGGTTTTCGCTCCACCAAGAGAGAGCGGAGCGCACTGCGCTTA 7109  
Qy 1085 TCGTGTCCAGACTTGGGGGTCCGTGTGTGCGAGAAATGGCCCTCTATGACGTGTCT 1144

Db 7110 TCGTATTCCAGATTGGGGGCTGCTGTGTGCGAGAAATAGCCCTTTAGAGTGTCT 7169  
QY 1145 CCAACCTCCCTCAGGCTGTGATGGGCTCTGTACGAGATTCCAGATTCTCTCGACAGC 1204  
Db 7170 CCAACCTCCCTCAGGCTGTGATGGGCTCTGTACGAGATTCTCTCGACAGC 7229  
QY 1205 GGGTCAGTTCCTGTGAAAGGCTGAAATCAAGAAAGACCCCTATGGGCTTTCATAG 1264  
Db 7230 GGGTCAGTTCCTGTGAAAGGCTGAAATCAAGAAAGACCCCTATGGGCTTTCATAG 7289  
QY 1265 ACAACCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTGAGAGATCAATT 1324  
Db 7290 ACAACCGCTGTTTGAATCAAGTCACTGAGATGACATCCGTGAGAGATCAATT 7349  
QY 1325 ATGATGTTGTGATCTGGCCCGGAAAGCAGACAGGCTATAGGCTGCTCAGAGGCGC 1384  
Db 7350 ACCAATGTTGTGATCTGGCCCGGAAAGCAGACAGGCTATAGGCTGCTCAGAGGCGC 7409  
QY 1385 TTTATATCGGGGCTCCCTGATCTAATTCAAAGGCGAATCGGCTATCGCCGCTGCC 1444  
Db 7410 TTTATATCGGGGCTCCCTGATCTAATTCAAAGGCGAATCGGCTATCGCCGCTGCC 7469  
QY 1445 GCGGAGCGGCTGTGAGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1504  
Db 7470 GCGGAGCGGCTGTGAGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 7529  
QY 1505 CTGACAGCTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1564  
Db 7530 CTGACAGCTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 7589  
QY 1565 TTGTCCTTATCTGTGAGAGGCGGGAACCTAGAGAGAGCGGAGGCTTACAGTCTTCA 1624  
Db 7590 TTGTCCTTATCTGTGAGAGGCGGGAACCTAGAGAGAGCGGAGGCTTACAGTCTTCA 7649  
QY 1625 CCGAGGCTATGACTAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1684  
Db 7650 CCGAGGCTATGACTAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 7709  
QY 1685 TGGAGTTGATACATCATGCTCTCCATGTGTGCTGCGGAGAGTATCTGGCAAAA 1744  
Db 7710 TGGAGTTGATACATCATGCTCTCCATGTGTGCTGCGGAGAGTATCTGGCAAAA 7769  
QY 1745 GGGTATACCTACCTCAGCCGAGACC 1769  
Db 7770 GGGTATACCTACCTCAGCCGAGACC 7794

RESULT 6  
US-09-539-601-10  
Sequence 10, Application US/09539601C  
Patent No. 6630343  
GENERAL INFORMATION:  
APPLICANT: Bartschlagel, Ralf FW  
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
FILE REFERENCE: all sequences  
CURRENT APPLICATION NUMBER: US/09/539,601C  
EARLIER FILING DATE: 2001-08-30  
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
EARLIER FILING DATE: 1999-04-03  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 10  
LENGTH: 7989  
TYPE: DNA  
ORGANISM: Hepatitis C virus  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: (1)..(341)  
OTHER INFORMATION: construct I377/NS3-3'/wt  
NAME/KEY: CDS  
LOCATION: (342)..(1181)

OTHER INFORMATION: hepatitis C virus core-neomycin phosphotransferase  
OTHER INFORMATION: fusion protein  
FEATURE:  
NAME/KEY: RBS  
LOCATION: (1190)..(1800)  
OTHER INFORMATION: internal ribosome entry site from  
OTHER INFORMATION: encephalomyocarditis virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1801)..(7758)  
OTHER INFORMATION: hepatitis C virus NS3 - 5B  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: (7759)..(7989)  
PUBLICATION INFORMATION:  
AUTHORS: Lohmann, Volker  
AUTHORS: Krieger, Frank  
AUTHORS: Koch, Jan-Oliver  
AUTHORS: Heitman, Ulrike  
AUTHORS: Theilmann, Lorenz  
AUTHORS: Bartschlagel, Ralf  
TITLE: Replication of subgenomic hepatitis C virus RNAs in a  
JOURNAL: Science  
VOLUME: 285  
PAGES: 110-113  
DATE: 1999-07-02  
US-09-539-601-10

Query Match 87.7%; Score 1552.2; DB 4; Length 7989;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1652; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 5 ATGACTTCCAGACCGCTGACCTGATGAGAGCCAACTCTGTGCGGATGAGATGGCG 64  
Db 5381 ATGACTTCCAGACCGCTGACCTGATGAGAGCCAACTCTGTGCGGATGAGATGGCG 5440  
QY 65 GGGACATTACCGCGTGTGAGTCAAGAAACAAGTATCTCTGAGCTCTTTCAGACCCG 124  
Db 5441 GGGACATTACCGCGTGTGAGTCAAGAAACAAGTATCTCTGAGCTCTTTCAGACCCG 5500  
QY 125 TCCGAGCGAGAGAGATGAGCGGAGATGCTCCGCGGAGATCTGCGGAATTCGA 184  
Db 5501 TCCGAGCGAGAGAGATGAGCGGAGATGCTCCGCGGAGATCTGCGGAATTCGA 5560  
QY 185 AGAAATTCACACGAGATGCGGATGAGGAGAGCGCGGATTCACACCTCCGCTGCG 244  
Db 5561 AGAAATTCACACGAGATGCGGATGAGGAGAGCGCGGATTCACACCTCCGCTGCG 5620  
QY 245 AGTCTGGAAGAGCGCGGATGAGTCTCTCCAGTGTATAGGTGCTCCAGTCCCACTTA 304  
Db 5621 AGTCTGGAAGAGCGCGGATGAGTCTCTCCAGTGTATAGGTGCTCCAGTCCCACTTA 5680  
QY 305 CTAAGACCCCTCTATACACCTCCAGGAGAAAGAGCAGTGTCTGACAGATTCGA 364  
Db 5681 CTAAGACCCCTCTATACACCTCCAGGAGAAAGAGCAGTGTCTGACAGATTCGA 5740  
QY 365 CCGTGTCTTCTGCGCTGAGGAGCTTGCACAAAGGCTTTTGTAGTCTCCGACCGTGG 424  
Db 5741 CCGTGTCTTCTGCGCTGAGGAGCTTGCACAAAGGCTTTTGTAGTCTCCGACCGTGG 5800  
QY 425 CCGTGAACGCGGACGAGCAACCGCCCTCTGACCAATCTTCCAGCAGCGGAGCAG 484  
Db 5801 CCGTGAACGCGGACGAGCAACCGCCCTCTGACCAATCTTCCAGCAGCGGAGCAG 5860  
QY 485 GATCTGAGTGTAGTGTATCTCTCCAGTCCCGCTTGAAGGAGGAGCGGAGGAGCCCG 544  
Db 5861 GATCTGAGTGTAGTGTATCTCTCCAGTCCCGCTTGAAGGAGGAGCGGAGGAGCCCG 5920  
QY 545 ATCTGAGCAGCGGCTTGTGTCTACCGTGTGAGAGAGCGGAGTGTGAGCGTGTGTCT 604  
Db 5921 ATCTGAGCAGCGGCTTGTGTCTACCGTGTGAGAGAGCGGAGTGTGAGCGTGTGTCT 5980

QY 605 GCTCGATGTCCTACATGAGCAGGGCGCTCTGATCAAGCCATGCGCTGCGAGGAAGCA 664  
Db 5981 GCTCGATGTCCTACATGAGCAGGGCGCTCTGATCAAGCCATGCGCTGCGAGGAAGCA 6040  
QY 665 AGCTGCCATCAACGGGTTGAGCACTTTGCTGCTGTCACCAACAATGTTCTGCTCA 724  
Db 6041 AGCTGCCATCAACGGGTTGAGCACTTTGCTGCTGTCACCAACAATGTTCTGCTCA 6100  
QY 725 CCACATCCCGGAGCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784  
Db 6101 CCACATCTTCGAGCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6160  
QY 785 TGGAGATCACTACAG 844  
Db 6161 TGGAGAGACACTACAG 6220  
QY 845 CTAACTCTTATCATGAG 904  
Db 6221 CTAACTCTTATCATGAG 6280  
QY 905 AATTGGCTATGAGGAG 964  
Db 6281 AATTGGCTATGAGGAG 6340  
QY 965 GCTCCGCTGAG 1024  
Db 6341 GCTCCGCTGAG 6400  
QY 1025 CAAAAAATGAGGTTTCTGCGCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084  
Db 6401 CAAAAAATGAGGTTTCTGCGCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6460  
QY 1085 TCGTGTCCCAAGCTTGGGGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1144  
Db 6461 TCGTGTCCCAAGCTTGGGGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6520  
QY 1145 CCACCTCCCTCAGGCGCTGATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1204  
Db 6521 CCACCTCCCTCAGGCGCTGATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6580  
QY 1205 GGGTGAAGTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1264  
Db 6581 GGGTGAAGTCTCTGCTGAG 6640  
QY 1265 ACACCCGCTGTTTGACTCAAGAGTCACTGAGATGAGATGAGATGAGATGAGATGAG 1324  
Db 6641 ACACCCGCTGTTTGACTCAAGAGTCACTGAGATGAGATGAGATGAGATGAGATGAG 6700  
QY 1325 ATCAATGTTGTAAGTGGGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1384  
Db 6701 ACCAATGTTGTAAGTGGGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6760  
QY 1385 TTTTATATGAGGAGTCCCTGATTAATTTAAAAAGGAGAGAGAGAGAGAGAGAGAGAG 1444  
Db 6761 TTTTATATGAGGAGTCCCTGATTAATTTAAAAAGGAGAGAGAGAGAGAGAGAGAGAG 6820  
QY 1445 GGGGAGAGCGGCTGCTGAG 1504  
Db 6821 GGGGAGAGCGGCTGCTGAG 6880  
QY 1505 CTGAGAGCTGTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1564  
Db 6881 CTGAGAGCTGTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6940  
QY 1565 TTTGTGTTATCTGTGAG 1624  
Db 6941 TTTGTGTTATCTGTGAG 7000  
QY 1625 CGGAGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 1684  
Db 7001 CGGAGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 7060  
QY 1685 TGGAGTTGATTAATCATGAGTCTCCATATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1744

Db 7061 TGGAGTTGATTAATCATGAGTCTCCATATGTTGTTGTTGTTGTTGTTGTTGTTGTT 7120  
QY 1745 GGGTATATACCTGACCCGTTGACCC 1769  
Db 7121 GGGTATATATCTGACCCGTTGACCC 7145  
  
RESULT 7  
US-09-539-601-7  
; Sequence 7, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FM  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; EARLIER APPLICATION NUMBER: 2001-08-30  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 8001  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct 1389/NS3-3'/wt  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1193)  
; OTHER INFORMATION: hepatitis C virus core-neomycin  
; OTHER INFORMATION: phosphotransferase fusion protein  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (1202)..(1812)  
; OTHER INFORMATION: internal ribosome entry site from  
; OTHER INFORMATION: encephalomyocarditis virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1813)..(7770)  
; OTHER INFORMATION: hepatitis C virus nonstructural proteins NS3-5B  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: (7771)..(8001)  
; PUBLICATION INFORMATION:  
; AUTHORS: Lohmann, Volker  
; AUTHORS: Krner, Frank  
; AUTHORS: Koch, Jan-Oliver  
; AUTHORS: Herian, Ulrike  
; AUTHORS: Theilmann, Lorenz  
; AUTHORS: Bartenschlager, Ralf  
; TITLE: Replication of subgenomic hepatitis C virus RNAs in a  
; TITLE: hepatoma cell line  
; JOURNAL: Science  
; VOLUME: 285  
; PAGES: 110-113  
; DATE: 1999-07-02  
; US-09-539-601-7  
  
Query Match 87.7%; Score 1552.2; DB 4; Length 8001;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;  
  
QY 5 ATGACTTCCAGAGCTGAGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 64  
Db 5393 ATGACTTCCAGAGAGCTGAGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5452  
QY 65 GGGACATTAACCCGGGTGAG 124  
Db 5453 GGAACATCAACCGGTGAG 5512

125 TCCGACGCGAGAGATGAGCCGGAAGTCTCCCGCGGAGATCTTCCGGAATCCA 184  
185 AGAATTTCCACAGAGATGCGCGATGAGGACGCGGATTTACAAACCTCCGCTCTGG 244  
5573 GGAATTTCCCTGAGAGATGCGCGATGAGGACGCGGATTTACAAACCTCCGCTCTGG 5632  
245 AGTCTGGAAGGCGCGGATGAGGACGCGGATTTACAAACCTCCGCTCTGG 304  
5633 AGTCTGGAAGGCGCGGATGAGGACGCGGATTTACAAACCTCCGCTCTGG 5692  
305 CTAAAGACCTCTCTATACCACTCTCCAGAGAAAGAGAGATTTGTTTCAAGATCCA 364  
5693 CCAAGGCGCGCTCGAATACCACTCTCCAGAGAAAGAGAGATTTGTTTCAAGATCCA 5752  
365 CCGTCTCTTCTGCGCGGAGCTTGCACAAAGCTTTTGGTAGCTCCGAGACCGTCCG 424  
5753 CCGTCTCTTCTGCGCGGAGCTTGCACAAAGCTTTTGGTAGCTCCGAGACCGTCCG 5812  
425 CCGTCTCTTCTGCGCGGAGCTTGCACAAAGCTTTTGGTAGCTCCGAGACCGTCCG 484  
5813 CCGTCTCTTCTGCGCGGAGCTTGCACAAAGCTTTTGGTAGCTCCGAGACCGTCCG 5872  
485 GATCTGACGTTGAGTGTATTTCTCCATGCCCCCTTGAAGGAGAGCCGCGGAGACCC 544  
5873 GATCTGACGTTGAGTGTATTTCTCCATGCCCCCTTGAAGGAGAGCCGCGGAGACCC 5932  
545 ATCTGACGCGGAGTGTATTTCTCCATGCCCCCTTGAAGGAGAGCCGCGGAGACCC 604  
5933 ATCTGACGCGGAGTGTATTTCTCCATGCCCCCTTGAAGGAGAGCCGCGGAGACCC 5992  
605 GCTGATGTTCTACATGAGCAAGGCGCTCTGATTCAGCCATGCGCTGCGAGAGAAAGCA 664  
5993 GCTGATGTTCTACATGAGCAAGGCGCTCTGATTCAGCCATGCGCTGCGAGAGAAAGCA 6052  
665 AGTCTGACGCGGAGTGTATTTCTCCATGCCCCCTTGAAGGAGAGCCGCGGAGACCC 724  
6053 AGTCTGACGCGGAGTGTATTTCTCCATGCCCCCTTGAAGGAGAGCCGCGGAGACCC 6112  
725 CCAATCTCCGCGGAGGAGCAAGCGGAGAGAAAGGTCACCTTGAAGAGCTGCAATCC 784  
6113 CCAATCTCCGCGGAGGAGCAAGCGGAGAGAAAGGTCACCTTGAAGAGCTGCAATCC 6172  
785 TGAGACATCACTACAGAGAGCTCTCAAGAGATGAGAGCGGAGCGTCCAGATTAGG 844  
6173 TGAGACATCACTACAGAGAGCTCTCAAGAGATGAGAGCGGAGCGTCCAGATTAGG 6232  
845 CTAAAGCTTCTATACATGAGAGAAAGCTGCAAGCTGACGCGGAGCGTCCAGATTAGG 904  
6233 CTAAAGCTTCTATACAGAGAGCTCTCAAGAGATGAGAGCGGAGCGTCCAGATTAGG 6292  
905 AATTGGCTATGAGGAGAGAGAGCTCCGGAACCTATCCAGAGCGGAGCGTCCAGATTAGG 964  
6293 AATTGGCTATGAGGAGAGAGAGCTCCGGAACCTATCCAGAGCGGAGCGTCCAGATTAGG 6352  
965 GCTCCGTTGAGAGAGAGCTTGTGGAAGACATGAAACCAATTTGACACCAATCATGAG 1024  
6353 GCTCCGTTGAGAGAGAGCTTGTGGAAGACATGAAACCAATTTGACACCAATCATGAG 6412  
1025 CAAAAAATGAGGTTTCTGCGTCCACCAAGAGAGAGCGCGGAGCGGAGCGTCCAGATTAGG 1084  
6413 CAAAAAATGAGGTTTCTGCGTCCACCAAGAGAGAGCGCGGAGCGGAGCGTCCAGATTAGG 6472  
1085 TCGTATTTCCAGAGATTTGAGGAGTGTGTCGAGAGAAAGGCGGAGCGGAGCGTCCAGATTAGG 1144  
6473 TCGTATTTCCAGAGATTTGAGGAGTGTGTCGAGAGAAAGGCGGAGCGGAGCGTCCAGATTAGG 6532  
1145 CCAACCTCTCTCAAGAGCTGTATGAGGAGCTCTGATACGAGATTCAGATTTCTCTGAGAGC 1204  
6533 CCAACCTCTCTCAAGAGCTGTATGAGGAGCTCTGATACGAGATTCAGATTTCTCTGAGAGC 6592

1205 GGGTGAAGTTCCTGTAAGAGCGCTGGAATCAAAAGAGACCTTATGAGGCTTTCATATG 1264  
6593 GGGTGAAGTTCCTGTAAGAGCGCTGGAATCAAAAGAGACCTTATGAGGCTTTCATATG 6652  
1265 ACAACCGGCTGTTTGAAGTCAAGAGTCACTGAGATGACATCCGTTGAGAGAGTCAATTT 1324  
6653 ACAACCGGCTGTTTGAAGTCAAGAGTCACTGAGATGACATCCGTTGAGAGAGTCAATTT 6712  
1325 ATCAATGTTGATGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1384  
6713 ACCAATGTTGATGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6772  
1385 TTTATATCGGAGGAGTCCCTGAGTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1444  
6773 TTTATATCGGAGGAGTCCCTGAGTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6832  
1445 GCGGAGCGGAGTCTGTAAG 1504  
6833 GCGGAGCGGAGTCTGTAAG 6892  
1505 CTGAGAGCTGTGAGAGTCAAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1564  
6893 CTGAGAGCTGTGAGAGTCAAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6952  
1565 TTGTGTTATCTGTGAG 1624  
6953 TTGTGTTATCTGTGAG 7012  
1625 CCGAGGCTATGAGTCAAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1684  
7013 CCGAGGCTATGAGTCAAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7072  
1685 TGAGGTTGATACATGATGCTCTCCATGATGCTGCTGCGAGAGAGAGAGAGAGAGAG 1744  
7073 TGAGGTTGATACATGATGCTCTCCATGATGCTGCTGCGAGAGAGAGAGAGAGAGAG 7132  
1745 GGGTATATACCTACCGGAG 1769  
7133 GGGTATATACCTACCGGAG 7157

RESULT 8  
US-09-539-601-4  
Sequence 4, Application US/09539601C  
Patent No. 6630343  
GENERAL INFORMATION:  
APPLICANT: Bartenschlager, Ralf FW  
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
FILE REFERENCE: all sequences  
CURRENT APPLICATION NUMBER: US/09/539,601C  
EARLIER FILING DATE: 2001-08-30  
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
NUMBER OF SEQ ID NOS: 51  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 8637  
TYPE: DNA  
ORGANISM: Hepatitis C virus  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: (1)..(341)  
OTHER INFORMATION: construct I377/NS2-3'/'wt  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (342)..(1181)  
OTHER INFORMATION: HCV core-neomycin phosphotransferase fusion  
OTHER INFORMATION: protein  
FEATURE:  
NAME/KEY: RBS  
LOCATION: (1190)..(1800)  
OTHER INFORMATION: internal ribosome entry site from  
OTHER INFORMATION: encephalomyocarditis virus

```

NAME/KEY: CDS
LOCATION: (1801)..(8406)
OTHER INFORMATION: hepatitis C virus NS2 - 5B
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (8407)..(8637)
PUBLICATION INFORMATION:
AUTHORS: Lohmann, Volker
AUTHORS: Krner, Frank
AUTHORS: Koch, Jan-Oliver
AUTHORS: Herian, Ulrike
AUTHORS: Theilmann, Lorenz
AUTHORS: Bartenschlager, Ralf
TITLE: Replication of subgenomic hepatitis C virus RNAs in a
JOURNAL: Science
VOLUME: 285
PAGES: 110-113
DATE: 1999-07-02
US-09-539-601-4

```

Query Match 87.7%; Score 1552.2; DB 4; Length 8637;

Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY	5	ATGACTTCCGAGACGCTACCTCATCGAGGGCAACTCTCTGTGGCGGATAGATAGGGCG	64
Db	6029	ATGACTCCCCGAGACCTACTCATCGAGGCCAACCTCTGTGGCGGAGGATAGGGCG	6086
QY	65	GGGACATTACCCGCGTGGAGTTCAGAGAA CAAAGTAGTAATCTTGGACTCTTTCGACCCGC	124
Db	6089	GGAACTATCACCCGCGTGGAGTTCAGAAATTAAGTAGTAATTTTGGACTCTTTCGAGCCGC	6148
QY	125	TTCCGAGCGGAGAGGATGATGCGGGAGAGTCCGTCCCGCGGAGATCTCTGGCGAAATCCA	184
Db	6149	TTCCAGCGGAGAGGATGATGAGGAGAGTATCCGTTCGCGGAGATCTCTGGAGATCCA	6208
QY	185	AGAAATTTCCACACAGCGATGCGCGCATGCGGACGCGCCGATTAACAACCTCCGCTGCTGG	244
Db	6209	GGAAATTTCTTCGAGCGATGCGCGCATATGGGACGCGCCGATTAACAACCTCCACCTGTTAG	6266
QY	245	AGTCTCGGAAGAGCCCCCGGACTACGTCTCTCCAGTGTATCATNGGTGTGCCACTGCGCACTTA	304
Db	6269	AGTCTCGGAAGAACCCCGGACTACGTCTCTCCAGTGTATCAGGGTGTCCATTTGCGCGCTG	6328
QY	305	CTAAGACCCCTCTATACCACTCCACGAGAGAAAGGACAGTTTTCATGACAGAAATCCA	364
Db	6329	CCAAGGCCCTCTCGATATCACCTCCACGAGAGAAAGGACAGTTTCTCTGTACAAATCTTA	6386
QY	365	CCGTGTCTTTCGCCCCGTGGCGAGCTTGCCACAAAGCTTTTGTGATCTCCGAGCCGTGG	424
Db	6389	CCGTGTCTTTCGCCCCGTGGCGAGCTTGCCACAAAGCTTTGCGGACGTCCGAATGTGG	6448
QY	425	CCGTGACAGGGGGCACCGGCAACCGCCCCCTCCGATACCAATCCCTCGACGACCGCGGAGAG	484
Db	6449	CCGTGACAGGGGGCACCGGCAACGGCTCTCTCTGACCAACCTTCGACACCGGACGCGGG	6508
QY	485	GATCTGACGTTGAGTGTATTTCTTCATGACCCCCCTTGAGGGGGAGCGGGGGACCCCCG	544
Db	6509	GATCCGACGTTGAGTGTATCTCTTCATGACCCCCCTTGAGGGGGAGCGGGGGATCTCCG	6566
QY	545	ATCTCAGGACGGGCTTTGTGTCTACCGTGAATGAGAGGGCCGGTAGAGACGTCTGTCTCT	604
Db	6569	ATCTCAGGACGGGCTTTGTGTCTACCGTGAACGAGAGAGCTTAGTAGACGTCTGTCTCT	6628
QY	605	GCTCGATGTCTTACACATGACAGGCGCTGTGATACCCCAATGCGCTCGGAGGAGAAACA	664
Db	6629	GCTCGATGTCTTACACATGACAGGCGCTGTGATACCCCAATGCCCTCGAGAGAGAAACA	6686
QY	665	AGCTGCCCATCAAGCGCTTAGAGCAACTCTTTGTCTGCTGACCAACAATGATCTTACGCTTA	724
Db	6689	AGCTGCCCATCAATGACATGAGCAACTCTTTGTCTCTGCTGACCAACAATGATCTTATGCTTA	6748

QY	725	CCACATCCCGAGCGCAAGCCAGCGGCGAGAGAAAGTCACTTTGACAGACTGCAATTC	784
Db	6749	CAACATCTCGAGCGCAAGCTCGGCGCGAAGAAAGTCACTTTGACAGACTGCGAGTCC	6808
QY	785	TGAGCAATCACTTACAGGACGTTGCTCAAGAGATGAAAGCGCAAGGCGCTCCACAGTAAAG	844
Db	6809	TGAGCACTTACCTACCGGACGTCGTCCAAAGAGATGAAGGCGCAAGGCGTCCACAGTAAAG	6866
QY	845	CTAAGCTTCTATCACTAGTAGAGAAAGCTCTCAAGCTGACGCCGCCCACTTCGGGCCAAATCTA	904
Db	6869	CTAATCTTCTATCCGTGAGAGAAAGCTCTTAAGCTGACGCCGCCCACTTCGGGCCAGATCTA	6922
QY	905	AATTTGCGTATGGGCGCAAAAGGACGTCCGAAACTATCCAGCAAGGCGCATTAACCATATCC	964
Db	6929	AATTTGGCTATGGGCGCAAAAGGACGTCGGAACCTATCCAGCAAGGCGGTTAAACCATATCC	6988
QY	965	GCTCCGTGTGGGAAGGACTTGTTTGGAAAGCATTGAAACCAATTGACACACATCATAGG	1022
Db	6989	GCTCCGTGTGGAAGGACTTGTCTGGAAAGCACTTGAGACACCAATTGACACCAATCATAGG	7048
QY	1025	CAAAAATAGAGTTTTCTGCCTCCAAACAGAGAGAGAGGCGCGCAAGCGAGCTGCCTTA	108
Db	7049	CAAAAATAGAGTTTTCTGCCTCCAAACAGAGAAAGGCGGCGCAAGCGAGCTGCCTTA	7108
QY	1085	TCGTATTCACGACTTGGGGGTCCGTGTGTGCGAAGAAATGAGCCCTCTATGACGTGTCT	1144
Db	7109	TCGTATTCACGAGATTTGGGGGTTGCTGTGTGTCGAGAAATAGGCCCTTTACGATGTGTCT	7168
QY	1145	CCACCTTCCTCAGGCTGTGATGGGCTCTCTGTACGATTCACAGTATTCCTTGACACGC	1204
Db	7169	CCACCTTCCTCAGGCGGTGATGGGCTCTTATACGGAATTCATATCTCTCTTGACACGC	7228
QY	1205	GAGTCGAGTTCCTGGTGAAGCGCTGGAATCAAGAAAGAACCCCTATGGGCTTTGCAATAG	1264
Db	7229	GAGTCGAGTTCCTGGTGAATCTCTGGAAAGCAAGAAATGCTCTTATGGGCTTTGCAATAG	7288
QY	1265	ACACCCGCTGTTTGACTCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGTCAATTT	1324
Db	7289	ACACCCGCTGTTTGACTCAACGGGCACTGAGAAATGACATCCGTGTAGAGAGTCAATCT	7348
QY	1325	ATCAATGTTTGACTTTGGCCCCCGAAGCGACACAGGCGATTAAGTTCGCTCACAGCGGC	1388
Db	7349	ACCAATGTTTGACTTTGGCCCCCGAAGCGACACAGGCGATTAAGTTCGCTCACAGCGGC	7408
QY	1385	TTTATATCGGGGGTCCCCCTGACTTAATCAAAAGGCGAAGCTGGGCGCTATGCGCGGTCC	1444
Db	7409	TTTACATCGGGGGCCCCCTGACTTAATTTAAAGGCGAAGCTGGGCGCTATGCGCGGTCC	7468
QY	1445	GCGCGAGCGGCGTCTGACGACTAGCTGCGGTAATACCTCTCACAATGTTACTTGAAGGCT	1504
Db	7469	GCGCGAGCGGCGTCTGACGACGACGCTGCGGTAATACCTCTCACAATGTTACTTGAAGGCG	7528
QY	1505	CTGCAGCCTGTGAGCTGCAAAAGCTCCAGACTGCACGATCGTGTGCGGAGACGCGC	1564
Db	7529	CTGCAGCCTGTGAGCTGCGAAGCTCCAGACTGCACGATCGTGTGCGGAGACGACAC	7588
QY	1565	TTTGTCGTTATCTGAGAGAGCGCGGGAAACCCAGAGAGACGCGCGAGACCTTAGAGTCTTCA	1624
Db	7589	TTTGTCGTTATCTGAGAAAGCGCGGGAGACCAAGAGAGACGAGGCGAGCTTACGGGCTTCA	7648
QY	1625	CGAGAGCTATGACTAGGTAATCTGCGCCCCCGCGGAGACCGCGCCCAACAGAAATACGAC	1684
Db	7649	CGAGAGCTATGACTAGATATCTTGTGCCCCCTGGGGACCGCGCCCAACAGAAATACGACT	7708
QY	1685	TGAGTTGATTAACATCAATGCTCTTCCAAATGTGTGCGTGCAGCAGATGCAATCTGGCAAAA	1744
Db	7709	TGAGTTGATTAACATCAATGCTCTTCCAAATGTGTGCAATGTGTGCGCAGATGCAATCTGGCAAAA	7766
QY	1745	GCGTATACACTCAACCGGTGACCC	1769
Db	7769	GCGTATACACTCAACCGGTGACCC	7793

```

RESULT 9
US-10-029-907-6
/ Sequence 6, Application US/10029907
/ Patent No. 6706874
/ GENERAL INFORMATION:
/ APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
/ TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
/ TITLE OF INVENTION: HEPATITIS C VIRUS
/ FILE REFERENCE: 13/083
/ CURRENT APPLICATION NUMBER: US/10/029,907
/ PRIOR FILING DATE: 2001-12-21
/ PRIOR APPLICATION NUMBER: 60/257,857
/ PRIOR FILING DATE: 2000-12-22
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6
/ LENGTH: 8638
/ TYPE: DNA
/ ORGANISM: HCV
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1802) ... (8407)
US-10-029-907-6

```

```

Query Match      87.7%; Score 1552.2; DB 4; Length 8638;
Blast Local Similarity 92.5%; Pred. No. 0;
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```

```

QY 5 ATGACTTCCGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGCATGAGATGGCG 64
DB 6030 ATGACTTCCGAGCGCTGACCTCATCGAGGCCAACCTCTGTGGCGCATGAGATGGCG 6089
QY 65 GGAACATTAACCCGCTGAGTCAAGAAACAAGTATCTCTGAGCTTTTGCACCCGC 124
DB 6090 GGAACATTAACCCGCTGAGTCAAGAAACAAGTATCTCTGAGCTTTTGCACCCGC 6149
QY 125 TCGACGCGAGAGAGATGAGCGGGAAGTGTCCGCGCGGAGATCTCCGGAATCCA 184
DB 6150 TCGACGCGAGAGAGATGAGCGGGAAGTGTCCGCGCGGAGATCTCCGGAATCCA 6209
QY 185 AGAATTTCCACAGCAGGATGCCCGCATGGGACGCGCGATTAACAACCTCCGCTGG 244
DB 6210 GGAATTTCCACAGCAGGATGCCCGCATGGGACGCGCGATTAACAACCTCCGCTGG 6269
QY 245 AGTCTGGAAGGCGCCGGAAGTACGTCCTTCAGTGTATCAATGGGTCCCACTGCCA 304
DB 6270 AGTCTGGAAGGCGCCGGAAGTACGTCCTTCAGTGTATCAATGGGTCCCACTGCCA 6329
QY 305 CTAAAGACCCCTCTATATCCACTCTCCACGAGAAAGAGACATGTTTCTGAAGATCCA 364
DB 6330 CTAAAGACCCCTCTATATCCACTCTCCACGAGAAAGAGACATGTTTCTGAAGATCCA 6389
QY 365 CCGTGTCTTCTGCGCTGAGCGAGCTTGCCACAAGGCTTTGGTAGCTCCGACCGTGG 424
DB 6390 CCGTGTCTTCTGCGCTGAGCGAGCTTGCCACAAGGCTTTGGTAGCTCCGACCGTGG 6449
QY 425 CCGTGAACAGCGGACGCGCAACGCGCTCTCTGAACCAATCTTCAGACGCGGAGCAG 484
DB 6450 CCGTGAACAGCGGACGCGCAACGCGCTCTCTGAACCAATCTTCAGACGCGGAGCAG 6509
QY 485 GATCTACGTTGATGATGATATCTCTCATGCGCCCTTGAAGGAGAGCGCGGAGACCCG 544
DB 6510 GATCTACGTTGATGATGATATCTCTCATGCGCCCTTGAAGGAGAGCGCGGAGACCCG 6569
QY 545 ATCTCAGCAGCGGCTTGTGCTTACCGTGAAGAGGCGCGGTGAAGACGTGCTGCT 604
DB 6570 ATCTCAGCAGCGGCTTGTGCTTACCGTGAAGAGGCGGTGAAGACGTGCTGCTGCT 6629
QY 605 GCTGATGCTTACATGAGACAGGCGCTTGAATCAAGCATGCGCTGCGAGAGAAAGCA 664
DB 6630 GCTGATGCTTACATGAGACAGGCGCTTGAATCAAGCATGCGCTGCGAGAGAAAGCA 6689

```

```

QY 665 AGTGCCTATCAACGCGTGTAGCAACTCTTGTGCTGCTACCAACAACATGCTTACGCTA 724
DB 6690 AGTGCCTATCAACGCGTGTAGCAACTCTTGTGCTGCTACCAACAACATGCTTACGCTA 6749
QY 725 CCACATCCCGACGCGCAACGCGGACAGAAAGGTCACTTTGACAGACTGCAATCC 784
DB 6750 CCACATCCCGACGCGCAACGCGGACAGAAAGGTCACTTTGACAGACTGCAAGTCC 6809
QY 785 TGAGCATCACTAACAGAGCGTGTCTAAGAGATGAAGGCGAAAGGCTCCACATTAAG 844
DB 6810 TGAGCATCACTAACAGAGCGTGTCTAAGAGATGAAGGCGAAAGGCTCCACATTAAG 6869
QY 845 CTAACTTCTATCAGTGAAGAGAGCGTGAAGTGAAGCGCCCACTTGGGCAATCTTA 904
DB 6870 CTAACTTCTATCAGTGAAGAGAGCGTGAAGTGAAGCGCCCACTTGGGCAATCTTA 6929
QY 905 AATTGGCTATGAGGCAAGAGAGCGTCCGAACTATCCAGAAAGCCATTAAACCATCC 964
DB 6930 AATTGGCTATGAGGCAAGAGAGCGTCCGAACTATCCAGAAAGCCATTAAACCATCC 6989
QY 965 GCTCCGTGTGAGAGACTTGTGGAAGACACTGAACACCAATTGACACCATCATATG 1024
DB 6990 GCTCCGTGTGAGAGACTTGTGGAAGACACTGAAGACACCAATTGACACCATCATATG 7049
QY 1025 CAAAAATAGAGTTTCTGCGTCCAAACGAGAGAGAGCGGCAAGCCAGCTCGCTTA 1084
DB 7050 CAAAAATAGAGTTTCTGCGTCCAAACGAGAGAGAGCGGCAAGCCAGCTCGCTTA 7109
QY 1085 TCGTGTTCAGACTTGTGGGGGTCCGTGTGCGAGAAATGGCCCTTATGACGTGTCT 1144
DB 7110 TCGTGTTCAGACTTGTGGGGGTCCGTGTGCGAGAAATGGCCCTTATGACGTGTCT 7169
QY 1145 CCACCTCCCTCAGAGTGTGATGGGTCTCTGTGAGCATTCAGATTTCTCTGACAC 1204
DB 7170 CCACCTCCCTCAGAGTGTGATGGGTCTCTGTGAGCATTCAGATTTCTCTGACAC 7229
QY 1205 GGGTGAAGTCTCGGTGAACGCGCTGGAATCAAGAAAGACCCCTATGAGCTTGTGATATG 1264
DB 7230 GGGTGAAGTCTCGGTGAACGCGCTGGAATCAAGAAAGACCCCTATGAGCTTGTGATATG 7289
QY 1265 ACACCGCTGTGTTTGACTCAACAGTCACTGAATATGACATCGGTGAGAGATCAATTT 1324
DB 7290 ACACCGCTGTGTTTGACTCAACAGTCACTGAATATGACATCGGTGAGAGATCAATTT 7349
QY 1325 ATCAATGTTGTGACTTGGGCCCCGAAAGCCAGACAGGCTTAAGGTGCTCAAGAGCGG 1384
DB 7350 ACCAATGTTGTGACTTGGGCCCCGAAAGCCAGACAGGCTTAAGGTGCTCAAGAGCGG 7409
QY 1385 TTTATATGAGGGGTCCTGACTAATTCAAAAGGCGAGAACTGCGGCTATGCGCGGTGC 1444
DB 7410 TTTATATGAGGGGTCCTGACTAATTCTAAAGGCGAGAACTGCGGCTATGCGCGGTGC 7469
QY 1445 GCGGAGCGGCTGTGCTGACGACTAGTGGGTATATCCCTCAATTTACTTTGAAGGCT 1504
DB 7470 GCGGAGCGGCTGTGCTGACGACTAGTGGGTATATCCCTCAATTTACTTTGAAGGCT 7529
QY 1505 CTGACGCTGTGAGTGTGCAAGGCTCAAGACTGACAGATGCTCGTGTGCGAGAGCGCC 1564
DB 7530 CTGCGGCTGTGAGTGTGCAAGGCTCAAGACTGACAGATGCTCGTGTGCGAGAGCGCC 7589
QY 1565 TTGTGTTATCTGTGAGAGCGGAGAACCAAGAGAGAGCGGCGAGCTTACAGAGTCTTA 1624
DB 7590 TTGTGTTATCTGTGAGAGCGGAGAACCAAGAGAGAGCGGCGAGCTTACAGAGTCTTA 7649
QY 1625 CCGAGGCTATGACTAGTACTTGTGCCCCCGGAGGACCCGCGCCAAACAGAAATAGACC 1684
DB 7650 CCGAGGCTATGACTAGTACTTGTGCCCCCGGAGGACCCGCGCCAAACAGAAATAGACC 7709
QY 1685 TGAAGTTGATACATCATCTCTCTCAATGTGTGCGTGGGAGAGATGATCTGGCAAAA 1744
DB 7710 TGAAGTTGATACATCATCTCTCTCAATGTGTGCGGAGAGATGATCTGGCAAAA 7769
QY 1745 GGGTATATACCTCAACCGGTGACCC 1769

```



Db 7770 GGGTGACTATCTCACCCGCTGACCC 7794  
RESULT 10  
US-10-029-907-24  
; Sequence 24, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802) ... (8407)  
US-10-029-907-24

Query Match 87.7%; Score 1552.2; DB 4; Length 8638;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 5 ATGACTTCCCAAGCGTGAAGCTGATGAGGCAACCTCTGTGCGGATGAGATGAGG 64  
Db 6030 ATGACTCCCCGAGACGTGACTCATGAGGCAACCTCTGTGCGGAGAGATGAGG 6089  
QY 65 GGAACATTACCCGCTGAGTCAAGAACAGGTAGTAATCTGCACTTTTCAGCCGC 124  
Db 6090 GGAACATCACCCGCTGAGTCAAGAAATTAAGTAGTAATTTTGACTTTTCAGCCGC 6149  
QY 125 TCCGAGCGGAGAGATGAGCGGGAAGTGTCCGTCCCGCGGAGATCTTCGCGAAATCCA 184  
Db 6150 TCCAGACGAGAGAGATGAGGAGAGATCCGTTCCGCGGAGATCTTCGCGAGATCCA 6209  
QY 185 AGAATATCCCAAGCGTGAAGCTGAGGCAAGCCCGGATTAACAACCTCTGCTGTG 244  
Db 6210 GGAATATCCCTGAGGAGATGCCATATGAGGCAAGCCCGATTAACAACCTCTGCTGTG 6269  
QY 245 AGTCTGGAAGGCCCGGACTAGTCCCTCCAGTGTAGATGAGTCCCACTGCCCTTA 304  
Db 6270 AGTCTGGAAGGCCCGGACTAGTCCCTCCAGTGTAGATGAGTGTCCATTGCCGCTG 6329  
QY 305 CTAAAGCCCCCTCTATACCACTCCACGGAAGAAAGGACATGTGTCTGAAGAAATCCA 364  
Db 6330 CCAAGGCCCTCTCGATACCACTCCACGGAAGAAAGGAGCGGTGTCTCTGATGAAATCCA 6389  
QY 365 CCGTGTCTTCTGCTGCGGAGCTTGCCAAAGGCTTTGTAGCTCCGAGCCGTGCG 424  
Db 6390 CCGTGTCTTCTGCTGCGGAGCTTGCCAAAGGCTTTGTAGCTCCGAGCTCCGAATGTGCG 6449  
QY 425 CCGTGAACGCGGACGCGAACCGCCCTCTGAGCAATCTCCAGCGAGCGGAGCAG 484  
Db 6450 CCGTGAACGCGGACGCGAACCGCCCTCTGAGCAATCTCCAGCGAGCGGAGCAGCAG 6509  
QY 485 GATTCGACGTTGAGTGTGATTTCTCGATGCCCCCTTTGAGGGGAGCGCGGGAGCCCG 544  
Db 6510 GATTCGACGTTGAGTGTGATTTCTCGATGCCCCCTTTGAGGGGAGCGCGGGAGATCCCG 6569  
QY 545 ATCTCAGCAGCGGCTTTGTGCTACCGTAGTGAAGGAGCGGTGAGAGCGTGTGCT 604  
Db 6570 ATCTCAGCAGCGGCTTTGTGCTACCGTAGTGAAGGAGCGGTGAGAGCGTGTGCT 6629  
QY 605 GCTCATGTCTTACATGAGACGCGCTGTGATCAAGCCATGCGCTGCGAGGAAAGCA 664

Db 6630 GCTCATGTCTTACATGAGACGCGCTGTGATCAAGCCATGCGCTGCGAGGAAACCA 6689  
QY 665 AGCTGCCATCAACGCCGTTGAGCAACTCTTTGCTGCGCTACCAACAATGTGTACGCTA 724  
Db 6690 AGCTGCCATCAATGACATGAGCAACTCTTTGCTGCGCTACCAACAATGTGTACGCTA 6749  
QY 725 CCACATCCCGAGCGCAAGCAGCGCGAGAAAGGTCACTTTGACAGACTGCAAAATCC 784  
Db 6750 CAACATCTCGAGCGCAAGCTCGCGGAGAAAGGTCACTTTGACAGACTGCAAGTCC 6809  
QY 785 TGAAGATCACTTACGAGAGCTGCTCAAGAGATGAAAGCGAAGCGCTCCAGTTAAG 844  
Db 6810 TGAAGATCACTTACGAGAGCTGCTCAAGAGATGAAAGCGAAGCGCTCCAGTTAAG 6869  
QY 845 CTAAGCTTCTATGAGTGAAGAGCGCTGAGCGCGCGCGCAATTCGCGCGCAATCTA 904  
Db 6870 CTAAGCTTCTATGAGTGAAGAGCGCTGAGCGCGCGCGCAATTCGCGCGCAATCTA 6929  
QY 905 AATTGGCTATGAGGCAAAAGACGTCCGGAACCTATCCAGCAAGGCCATTAAACCATCC 964  
Db 6930 AATTGGCTATGAGGCAAAAGACGTCCGGAACCTATCCAGCAAGGCCATTAAACCATCC 6989  
QY 965 GCTCGTGTGAGAGACTTGTGGAAGCACTGAACCAATTTGACACCAATCATG 1024  
Db 6990 GCTCGTGTGAGAGACTTGTGGAAGCACTGAAGCACTGAATTTGACACCAATCATG 7049  
QY 1025 CAAAAAATGAGTGTTCGCGTCCACAGAGAGAGGCGGCAAGCCAGCTGCGCTTA 1084  
Db 7050 CAAAAAATGAGTGTTCGCGTCCACAGAGAGAGGCGGCAAGCCAGCTGCGCTTA 7109  
QY 1085 TCGTGTTCACAGACTTGGGGTCCGTGTGTCGAGAAATGAGCCCTCATGACGTGTCT 1144  
Db 7110 TCGTATTCACGATTTTGGGGTGTGTGTGTCGAGAAATGAGCCCTTTACATGTGTCT 7169  
QY 1145 CCACCTCTCTGAGGCTGTGATGAGCTCTCTGTAACGATTTCAAGTATCTCTGACAGC 1204  
Db 7170 CCACCTCTCTGAGGCTGTGATGAGCTCTCTGTAACGATTTCAAGTATCTCTGACAGC 7229  
QY 1205 GGGTGAAGTCTGCTGTAAGCGCTGGAATCAAAAGGACCCCTATGAGGCTTCATATG 1264  
Db 7230 GGGTGAAGTCTGCTGTAAGCGCTGGAATCAAAAGGATGCTTCATGAGGCTTCATATG 7289  
QY 1265 AACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGAGAGAGTCAATTT 1324  
Db 7290 AACCCGCTGTTTGACTCAACAGTCACTGAGATGACATCCGTGAGAGAGTCAATTT 7349  
QY 1325 ATCAATGTTGACTTGAGCCCGGAGAGCAGGCGCATTAAGTCCGTCAAGAGCGGC 1384  
Db 7350 ATCAATGTTGACTTGAGCCCGGAGAGCAGGCGCATTAAGTCCGTCAAGAGCGGC 7409  
QY 1385 TTTATATGAGGAGTCCCTGCTAATTCAAAAGGAGCAATGAGGCTATGAGCGGTGCC 1444  
Db 7410 TTTATATGAGGAGTCCCTGCTAATTCAAAAGGAGCAATGAGGCTATGAGCGGTGCC 7469  
QY 1445 GCGGAGCGGAGTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1504  
Db 7470 GCGGAGCGGAGTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 7529  
QY 1505 CTGAGCCTGTGAGCTGCAAAAGCTCCAGACTGACAGATGCTGTGTGCGGAGACGCGC 1564  
Db 7530 CTGAGCCTGTGTGAGCTGCAAAAGCTCCAGACTGAGCTGAGCTGAGCTGAGCTGAGC 7589  
QY 1565 TTTGCTATCTGAGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACAGATCTTCA 1624  
Db 7590 TTTGCTATCTGAGAGAGCGCGGGAACCCAGAGAGAGCGCGGAGCTTACAGATCTTCA 7649  
QY 1625 CGAGAGCTATGACTAGTACTCTGCCCCCGGAGGACCCGCGCAACAGAAATACGACC 1684  
Db 7650 CGAGAGCTATGACTAGTACTCTGCCCCCGGAGGACCCGCGCAACAGAAATACGACT 7709  
QY 1685 TGAAGTTGATTAATCATGCTCTTCCAAATGTGTGCGGTGCGCAAGATCATCTGCAAAA 1744



Db 7710 TGGAGTTGATACATCATGCTCTCAATGTCAGTCGGGACGATGCATCTGGCAAAA 7769  
Qy 1745 GGGTATCTACCTCAGCCCGTGACC 1769  
Db 7770 GGGTATCTACCTCAGCCCGTGACC 7794

## RESULT 11

US-10-029-907-1  
Sequence 1, Application US/10029907  
Patent No. 6706874  
GENERAL INFORMATION:  
APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
HEPATITIS C VIRUS  
FILE REFERENCE: 13/083  
CURRENT APPLICATION NUMBER: US/10/029,907  
PRIOR FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/257,857  
PRIOR FILING DATE: 2000-12-22  
NUMBER OF SEQ. ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ. ID NO 1  
LENGTH: 8639  
TYPE: DNA  
ORGANISM: HCV  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1803) ... (8408)  
US-10-029-907-1

Query Match 87.7%; Score 1552.2; DB 4; Length 8639;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 5 ATGACCTCCGAGCGTGAAGCTCATGAGGCAACCTCTGAGCGGATGAGTGGCG 64  
Db 6031 ATATCTCCCGGAGCGTGAAGCTCATGAGGCAACCTCTGAGCGGATGAGTGGCG 6090  
Qy 65 GGGACATTACCCGCGTGAAGCTCATGAGGCAACCTCTGAGCTCTTTCAGCCGC 124  
Db 6091 GGAACATCAACCGCGTGAAGCTCATGAGGCAACCTCTGAGCTCTTTCAGCCGC 6150  
Qy 125 TCGAGCGGAGGAGTGAAGCGGGAAGTGTCCGTCGCGGAGATCTTCGGAATCCA 184  
Db 6151 TCGAGCGGAGGAGTGAAGCGGGAAGTGTCCGTCGCGGAGATCTTCGGAATCCA 6210  
Qy 185 AGAATTTCCAGCAGGATGCGGATGAGGCAACCTCTGAGCTCTTTCAGCCGC 244  
Db 6211 GGAATTTCCAGCAGGATGCGGATGAGGCAACCTCTGAGCTCTTTCAGCCGC 6270  
Qy 245 AGTCTGGAAGGCGCGGATGAGGCAACCTCTGAGCTCTTTCAGCCGC 304  
Db 6271 AGTCTGGAAGGCGCGGATGAGGCAACCTCTGAGCTCTTTCAGCCGC 6330  
Qy 305 CTAAGACCCCTCTTATACCACTTCAGGAGAAAGAGAGAGATTTTCTCAAGATCCA 364  
Db 6331 CTAAGACCCCTCTTATACCACTTCAGGAGAAAGAGAGAGATTTTCTCAAGATCCA 6390  
Qy 365 CCGTCTTCTGCGCGGAGGATTCGCAAAAGCTTTGAGTCTCGGACCGTGG 424  
Db 6391 CCGTCTTCTGCGCGGAGGATTCGCAAAAGCTTTGAGTCTCGGACCGTGG 6450  
Qy 425 CCGTCAAGCGGAGCGGCAACCGCTTCGCAACATCTTCGAGCAGCGGAGAG 484  
Db 6451 CCGTCAAGCGGAGCGGCAACCGCTTCGCAACATCTTCGAGCAGCGGAGAG 6510  
Qy 485 GATCTGACCTTGAAGTGTATTTCTCTCATGCCCCCTTGAAGGAGAGCGGAGAGCCCG 544  
Db 6511 GATCTGACCTTGAAGTGTATTTCTCTCATGCCCCCTTGAAGGAGAGCGGAGAGCCCG 6570  
Qy 545 ATCTGAGCGAGGAGCTTGGTCTACGATGAGGAGGCGGAGAGAGAGTGTCTGCT 604  
Db ATCTGAGCGAGGAGCTTGGTCTACGATGAGGAGGCGGAGAGAGAGTGTCTGCT 604

Db 6571 ATCTGAGCGAGGAGCTTGGTCTACGATGAGGAGGCGGAGAGAGTGTCTGCT 6630  
Qy 605 GCTGATGCTTACATAGGAGAGCGGCTCTGATTCAGCCATGCGCTCGGAGAGAGCA 664  
Db 6631 GCTGATGCTTACATAGGAGAGCGGCTCTGATTCAGCCATGCGCTCGGAGAGAGCA 6690  
Qy 665 AGCTGCCATCAACGCGTTGAGCACTTTGCTGCTGCAACCAATGCTTCAAGCTTA 724  
Db 6691 AGCTGCCATCAACGCGTTGAGCACTTTGCTGCTGCAACCAATGCTTCAAGCTTA 7300  
Qy 725 CCAATCCCGCAGCGAGCGGAGCGGAGAGAGTCACTTTGACAGACTGCAATCC 784  
Db 6751 CCAATCCCGCAGCGAGCGGAGCGGAGAGAGTCACTTTGACAGACTGCAATCC 6810  
Qy 785 TGAACGATCTACGAGAGCTGCTCAAGAGATGAAGCGGAGGCTTCAAGTAAAG 844  
Db 6811 TGAACGATCTACGAGAGCTGCTCAAGAGATGAAGCGGAGGCTTCAAGTAAAG 8870  
Qy 845 CTAAGCTTCTATAGTGAAGAGAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTTA 904  
Db 6871 CTAAGCTTCTATAGTGAAGAGAGCTGCAAGCTGCAAGCTGCAAGCTTCAAGCTTA 6930  
Qy 905 AATTGGCTATGAGGAGAGAGAGCTGCAAGCTTATCAAGAGCGGATTAACCATCC 964  
Db 6931 AATTGGCTATGAGGAGAGAGAGCTGCAAGCTTATCAAGAGCGGATTAACCATCC 6990  
Qy 965 GCTCGCTGAGAGAGCTTGTGAAGAGAGCTGCAAGAGAGAGAGAGAGAGAGAG 1024  
Db 6991 GCTCGCTGAGAGAGCTTGTGAAGAGAGCTGCAAGAGAGAGAGAGAGAGAGAG 7050  
Qy 1025 CAAAAAATGAGGTTTCTGCTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084  
Db 7051 CAAAAAATGAGGTTTCTGCTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7110  
Qy 1085 TCGTGTCCAGACTTGGGAGTCCGCTGTGTCAGAGAGAGAGAGAGAGAGAGAG 1144  
Db 7111 TCGTGTCCAGACTTGGGAGTCCGCTGTGTCAGAGAGAGAGAGAGAGAGAGAG 7170  
Qy 1145 CACCTCTCTCAGAGCTGATGAGGCTCTGCAAGAGATTCAGATTTCTCTGAGCAGC 1204  
Db 7171 CACCTCTCTCAGAGCTGATGAGGCTCTGCAAGAGATTCAGATTTCTCTGAGCAGC 7230  
Qy 1205 GGGTGAAGTCTGAGGAGAGAGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1264  
Db 7231 GGGTGAAGTCTGAGGAGAGAGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 7290  
Qy 1265 ACAACCGCTGTTTGAAGTCAAGAGTCAAGAGATCAAGAGAGAGAGAGAGAGAG 1324  
Db 7291 ACAACCGCTGTTTGAAGTCAAGAGTCAAGAGATCAAGAGAGAGAGAGAGAGAG 7350  
Qy 1325 ATCAATGTTGATCTTGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1384  
Db 7351 ATCAATGTTGATCTTGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7410  
Qy 1385 TTTATATGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1444  
Db 7411 TTTATATGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7470  
Qy 1445 GCGGAGAGGAGGAG 1504  
Db 7471 GCGGAGAGGAGGAG 7530  
Qy 1505 CTGAGAGCTGTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCA 1564  
Db 7531 CTGAGAGCTGTGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCA 7590  
Qy 1565 TTGCTGTTATCTGAG 1624  
Db 7591 TTGCTGTTATCTGAG 7650  
Qy 1625 CCGAGGCTATGAGTGAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1684  
Db 7651 CCGAGGCTATGAGTGAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7710

QY 1685 TGGAGTTGATTAACATCATGCTCTCCAAATGTGTGGTCCGCGACATGATCTGCAAAA 1744  
| | | | |  
Db 7711 TGGAGTTGATTAACATCATGCTCTCCAAATGTGTGGTCCGCGACATGATCTGCAAAA 7770  
| | | | |  
QY 1745 GGGTATACTACTCTACCCCGTACCC 1769  
| | | | |  
Db 7771 GGGTATACTACTCTACCCCGTACCC 7795  
| | | | |  
RESULT 12  
US-10-029-907-4  
; Sequence 4, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; HEPATITIS C VIRUS  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 8643  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802) ... (8407)  
US-10-029-907-4  
Query Match 87.7%; Score 1552.2; DB 4; Length 8643;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1632; Conservative 0; Mismatches 133; Indels 0; Gaps 0;  
QY 5 ATGACTTCCAGACGCTGATCTCATGAGCCCACTCTGTGGCGCATGAGATGAGCG 64  
| | | | |  
Db 6030 ATGACTTCCAGACGCTGATCTCATGAGCCCACTCTGTGGCGCATGAGATGAGCG 6089  
| | | | |  
QY 65 GGGACATTACCGCGCTGAGTCAAGAACAGGTAGTAATCTTGAATCTTTCGACCCGC 124  
| | | | |  
Db 6090 GGGACATTACCGCGCTGAGTCAAGAACAGGTAGTAATCTTGAATCTTTCGACCCGC 6149  
| | | | |  
QY 125 TCCGACCGGAGAGATGAGCGGAGATGCTCCGCGCGAGATCTCCGCGGAATCCA 184  
| | | | |  
Db 6150 TCCGACCGGAGAGATGAGCGGAGATGCTCCGCGCGAGATCTCCGCGAGTCCA 6209  
| | | | |  
QY 185 AGAAATTTCCACAGAGATGCGCATGAGGACAGCCCGATTCAACCTCCGCTGTGG 244  
| | | | |  
Db 6210 AGAAATTTCCACAGAGATGCGCATGAGGACAGCCCGATTCAACCTCCGCTGTGG 6269  
| | | | |  
QY 245 AGTCTGGAAGGCCCGGACATACTCTCCAGTGTATAGTGGTCCCACTGCCACTTA 304  
| | | | |  
Db 6270 AGTCTGGAAGGCCCGGACATACTCTCCAGTGTATAGTGGTCCCACTGCCACTTA 6329  
| | | | |  
QY 305 CTAAGACCCCTCTATACCACTCCACGGAAGAGAGACATTTGTTGACAGAAATCCA 364  
| | | | |  
Db 6330 CTAAGACCCCTCTATACCACTCCACGGAAGAGAGACATTTGTTGACAGAAATCCA 6389  
| | | | |  
QY 365 CCGTGTCTTCTGCTCTGCGGAGCTTGCACAAAGGCTTTGGTAGCTCCGAGCGTGG 424  
| | | | |  
Db 6390 CCGTGTCTTCTGCTCTGCGGAGCTTGCACAAAGGCTTTGGTAGCTCCGAGCTGG 6449  
| | | | |  
QY 425 CCGTGAACGCGGACAGCGCAACCGCCCTCTGACCAATCTTCCAGCGAGGAGCG 484  
| | | | |  
Db 6450 CCGTGAACGCGGACAGCGCAACCGCCCTCTGACCAATCTTCCAGCGAGGAGCG 6509  
| | | | |  
QY 485 GATTCGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544  
| | | | |  
Db 6510 GATTCGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6569  
| | | | |

QY 545 ATCTCAGGACAGGGGCTTGTGTCTACCTGATGAGAGAGCGGTGAGGACGTCGTGCT 604  
| | | | |  
Db 6570 ATCTCAGGACAGGGGCTTGTGTCTACCTGATGAGAGAGCGGTGAGGACGTCGTGCT 6629  
| | | | |  
QY 605 GCTGATGTCTTACACATGAGACAGGCGCTTGTATCAGCCCATGCGTCGAGAGAACCA 664  
| | | | |  
Db 6630 GCTGATGTCTTACACATGAGACAGGCGCTTGTATCAGCCCATGCGTCGAGAGAACCA 6689  
| | | | |  
QY 665 AGCTGCCATACACCGGTGAGCAACTTTTGTGCGCTCACCAACATGTGTACGCTA 724  
| | | | |  
Db 6690 AGCTGCCATACACCGGTGAGCAACTTTTGTGCGCTCACCAACATGTGTACGCTA 6749  
| | | | |  
QY 725 CCACATCCCGGACGCGCAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784  
| | | | |  
Db 6750 CCACATCCCGGACGCGCAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6809  
| | | | |  
QY 785 TGGACGATCACTACCGAGACGCTGCTCAAGAGATGAGAGAGAGAGAGAGAGAGAGAG 844  
| | | | |  
Db 6810 TGGACGATCACTACCGAGACGCTGCTCAAGAGATGAGAGAGAGAGAGAGAGAGAGAG 6869  
| | | | |  
QY 845 CTAACTTCTATCACTAG 904  
| | | | |  
Db 6870 CTAACTTCTATCACTAG 6929  
| | | | |  
QY 905 AATTGGCTATGAGGCAAG 964  
| | | | |  
Db 6930 AATTGGCTATGAGGCAAG 6989  
| | | | |  
QY 965 GCTCGGTGAG 1024  
| | | | |  
Db 6990 GCTCGGTGAG 7049  
| | | | |  
QY 1025 CAAAAAATGAGGTTTTCGCGTCCAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084  
| | | | |  
Db 7050 CAAAAAATGAGGTTTTCGCGTCCAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7109  
| | | | |  
QY 1085 TCGTGTCTCCAGACTTGGGAGTCCGTTGTGTGTCGAGAAATGAGGCTCTATGACGTGCT 1144  
| | | | |  
Db 7110 TCGTGTCTCCAGACTTGGGAGTCCGTTGTGTGTCGAGAAATGAGGCTCTATGACGTGCT 7169  
| | | | |  
QY 1145 CCACCTCTCTGAGGCTGTGATGAGGCTCTCTGTACGATTCGATGATCTCTCTGAGACGC 1204  
| | | | |  
Db 7170 CCACCTCTCTGAGGCTGTGATGAGGCTCTCTGTATACGATTCGATGATCTCTCTGAGACGC 7229  
| | | | |  
QY 1205 GGGTCAAGTCTCGGTGAGAGCGCTGAGAAATCAAGAGAACCCCTATGAGGCTTGTGATG 1264  
| | | | |  
Db 7230 GGGTCAAGTCTCGGTGAGAGCGCTGAGAAATCAAGAGAACCCCTATGAGGCTTGTGATG 7289  
| | | | |  
QY 1265 ACACCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGATCAATTT 1324  
| | | | |  
Db 7290 ACACCGCTGTTTGAATCAACAGTCACTGAGAAATGACATCCGTGTAGAGAGATCAATTT 7349  
| | | | |  
QY 1325 ATCAATGTTGACTTGGCCCCCGAGAGCGACAGGCCATAGAGTCTCTACAGAGCGGC 1384  
| | | | |  
Db 7350 ATCAATGTTGACTTGGCCCCCGAGAGCGACAGGCCATAGAGTCTCTACAGAGCGGC 7409  
| | | | |  
QY 1385 TTTAATGAGGAGGTCCTCTGACTAATTCAAAAGGCGAGAACTGAGGCTATGAGCGGTGCC 1444  
| | | | |  
Db 7410 TTTAATGAGGAGGTCCTCTGACTAATTCAAAAGGCGAGAACTGAGGCTATGAGCGGTGCC 7469  
| | | | |  
QY 1445 GCGGAGCGGAGTGTGAGTCACTAGCTGCGGTAAATTAACCTCACAATGTTACTTGAAGGCT 1504  
| | | | |  
Db 7470 GCGGAGCGGAGTGTGAGTCACTAGCTGCGGTAAATTAACCTCACAATGTTACTTGAAGGCT 7529  
| | | | |  
QY 1505 CTGACGCTGTGAGCTGCAAAAGCTCCAGAGCTGACGATGCTGTGTGCGAGAGCGGCC 1564  
| | | | |  
Db 7530 CTGACGCTGTGAGCTGCAAAAGCTCCAGAGCTGACGATGCTGTGTGCGAGAGCGACC 7589  
| | | | |  
QY 1565 TTTGCTTATCTGTGAGAGCGCGGAGAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1624  
| | | | |  
Db 7590 TTTGCTTATCTGTGAGAGCGCGGAGAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7649  
| | | | |





```

|||||
9248 TGGACACCACTACCGGAGCTGCTCAGAGAGTGAAGCGAGGTCACAGCTTAAG 9307
QY CTAGCTTTATCATGAGAGAAAGCTCGAAGCTGAGCCGCCCACTTGCGCCAATCTA 904
DB CTAAACTTCTATCCGAGAGAAAGCTGTAGCTGAGCCGCCCACTTGCGCCAATCTA 9367
QY AATTGGCTATGGGCGAAGAGAGCTCCGGAACCTATCCAGAGAGGCACTTAACCAATCC 964
DB AATTGGCTATGGGCGAAGAGAGCTCCGGAACCTATCCAGAGAGGCACTTAACCAATCC 9427
QY GCTCCGTGGAGAGAGCTTTGGAAGACATGAAACCAATTGACACCAACCATCATG 1024
DB GCTCCGTGGAGAGAGCTTTGGAAGACATGAAACCAATTGACACCAACCATCATG 9487
QY 1025 CAAAAAATGAGTTTCTCGCTGCAACAGAGAGAGGCGCAAGCAAGCTCGCTTA 1084
DB CAAAAAATGAGTTTCTCGCTGCAACAGAGAGAGGCGCAAGCAAGCTCGCTTA 9547
QY 1085 TCGTGTCCAGACTTGGGGGTCCGTGTGCGAGAAATGGCCCTCTATGAGTGTCT 1144
DB TCGTGTCCAGACTTGGGGGTCCGTGTGCGAGAAATGGCCCTCTATGAGTGTCT 9607
QY 1145 CCACCTCCCTCAGGCTGATGAGGCTCTCGTACGATTTCCAGTATCTCTCGAGAC 1204
DB CCACCTCCCTCAGGCTGATGAGGCTCTCGTACGATTTCCAGTATCTCTCGAGAC 9667
QY 1205 GGGTCAGTTCCTGTGTGAACGCTGTGAATCAAGAAAGACCCCTATGGGCTTTGATATG 1264
DB GGGTCAGTTCCTGTGTGAATGCTGTGAAGCAAGAAATGGCCCTATGGGCTTTGATATG 9727
QY 1265 AACCCGCTGTTTGACTCAAGTCACTGAGAAATGACATCCGTGTAGAGAGTCAATT 1324
DB AACCCGCTGTTTGACTCAAGTCACTGAGAAATGACATCCGTGTAGAGAGTCAATT 9787
QY 9728 AACCCGCTGTTTGACTCAAGTCACTGAGAAATGACATCCGTGTAGAGAGTCAATT 9787
DB 1325 ATCAATGTTGTGACTTGAGCCCGCGAAGCAGACAGGCACTTAAGTGTCTACAGAGCG 1384
QY 9788 ACBAATGTTGTGACTTGAGCCCGCGAAGCAGACAGGCACTTAAGTGTCTACAGAGCG 9847
DB 9848 TTTATATCGGGGGTCCCTGACTAATTTCAAAAGGAGAGACTGCGGCTATCGCGGTG 1444
QY 1445 GCGGAGCGGCTGCTGAGACTGAGTGGGTATATCCCTCAGATGTTACTTGAAGGCT 1504
DB GCGGAGCGGCTGCTGAGACTGAGTGGGTATATCCCTCAGATGTTACTTGAAGGCT 9967
QY 1505 CTGCAAGCTGTGAGAGCTGCAAGAGCTCAGAGAGTGCATGCTGCTGTGCGAGAGCG 1564
DB CTGCAAGCTGTGAGAGCTGCAAGAGCTCAGAGAGTGCATGCTGCTGTGCGAGAGCG 10027
QY 1565 TTGTCGTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGAGAGCTTACAGAGTCTTCA 1624
DB TTGTCGTTATCTGTGAGAGCGCGGGAACCCAGAGAGAGCGCGAGAGCTTACAGAGTCTTCA 10087
QY 1625 CGGAGGCTATGACTAGGTAATCTGCCCCCGCGGAGACCGGCCCAACCAAGTAATGAGCC 1684
DB CGGAGGCTATGACTAGGTAATCTGCCCCCGCGGAGACCGGCCCAACCAAGTAATGAGCC 10147
QY 1685 TGGAGTGTATACATGATGCTCTCAATGTGTGCGTGTGCGAGAGTGTGATCTGGGAGAA 1744
DB TGGAGTGTATACATGATGCTCTCAATGTGTGCGTGTGCGAGAGTGTGATCTGGGAGAA 10207
QY 1745 GGGTATACCTACACCGGTGAGCC 1769
DB GGGTATACCTACCTACACCGGTGAGCC 10232

```

RESULT 15  
 US-10-029-907-5  
 ; Sequence 5, Application US/10029907  
 ; Patent No. 6706874  
 ; GENERAL INFORMATION:

```

; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 8648
; TYPE: DNA
; ORGANISM: HCV
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
; US-10-029-907-5

```

```

Query Match      87.6%; Score 1550.6; DB 4; Length 8648;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 1631; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```

```

QY 5 ATGACTTCCAGAGCGCTGACTCATGAGGCGCAACTCCTGTGCGGAGATGAGAGCG 64
DB 6030 ATGACTTCCAGAGCGCTGACTCATGAGGCGCAACTCCTGTGCGGAGATGAGAGCG 6089
QY 65 GGGACATTAACCGCGTGTGAGTCAAGAAAGAGTATCTGTGAGCTTTGACCCGC 124
DB 6090 GGAACATCAACCGCGTGTGAGTCAAGAAATATGATGATTTTGGAGCTTTGAGCCGC 6149
QY 125 TCCGAGCGGAGAGATGAGCGGGAAGTCCGTCCCGCGGAGATCCCTGCGGAAATCA 184
DB 6150 TCCAGCGGAGAGAGATGAGCGGGAAGTCCGTCCCGCGGAGATCCCTGCGGAGATCA 6209
QY 185 AGAAATTCACAGAGATGAGCGGAGAGCGCGGAGATTAACAACCTCCGCTGCTG 244
DB 6210 GGAATTTCCCTGAGAGATGAGCGGAGAGCGCGGAGATTAACAACCTCCGCTGCTG 6269
QY 245 AGTCTGGAAGCGCGGAGATGAGCGGAGAGCGCGGAGATTAACAACCTCCGCTGCTG 304
DB 6270 AGTCTGGAAGCGCGGAGATGAGCGGAGAGCGCGGAGATTAACAACCTCCGCTGCTG 6329
QY 305 CTAAAGCCCTCTATACCACTTCAAGAGAAAGAGACATGTTGTTGAGAGATCA 364
DB 6330 CCAAGCCCTCTATACCACTTCAAGAGAAAGAGACATGTTGTTGAGAGATCA 6389
QY 365 CCGTGTCTTGTGCGCGGAGAGCTTGCACAAAGGCTTTGATGAGTCCGAGAGCTGCG 424
DB 6390 CCGTGTCTTGTGCGCGGAGAGCTTGCACAAAGGCTTTGATGAGTCCGAGAGCTGCG 6449
QY 425 CCGTGTGAGCGGAGAGCGGAGAGCGCGCTCTGAGCAATCTCTGAGAGAGAGAG 484
DB 6450 CCGTGTGAGCGGAGAGCGGAGAGCGCGCTCTGAGCAATCTCTGAGAGAGAGAG 6509
QY 485 GATCTGAGCTTGAATGATTTCTTCAATGCGCGGAGAGAGAGAGAGAGAGAGAG 544
DB 6510 GATCTGAGCTTGAATGATTTCTTCAATGCGCGGAGAGAGAGAGAGAGAGAGAG 6569
QY 545 ATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
DB 6570 ATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6629
QY 605 GCTGTGATGCTTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664
DB 6630 GCTGTGATGCTTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6689
QY 665 AGCTGCCATCAACGCGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
DB 6690 AGCTGCCATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6749
QY 725 CCAACATCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784

```

Db 6750 CAACATCTGCAGCCGCAACCTCTGCGCAGAGAAGGTCACTTTGACAGACTGCAGGTCC 6809  
QY 785 TGGAGATACATACACAGAGCGTGTCAAGAGATGAAGGCGTCCACAGTTAAAG 844  
Db 6810 TGGAGACCACTACCGGACGTCGCTCAAGAGATGAAGGCGTCCACAGTTAAAG 6869  
QY 845 CTAGCTTCTATCAGTAGAGGAAGCTGCAAGCTGACGCCCACTTGCGCAATCTA 904  
Db 6870 CTAACTTCTATCCGTGAGAGGAAGCTGTAAAGCTGACGCCCACTTGCGCAATCTA 6929  
QY 905 AATTGGCTATGGGCAAGAGAGTCGCGAACCTTATCCAGCAAGGCCATTAAACCATCC 964  
Db 6930 AATTGGCTATGGGCAAGAGAGTCGCGAACCTTATCCAGCAAGGCCGTTAAACCATCC 6989  
QY 965 GTCCTGTTGGAGAGACTTTGGAAAGACATGAAACCAATTGACACCAACCATCATGG 1024  
Db 6990 GTCCTGTTGGAGAGACTTTGGAGAGACTGAGACACCAATTGACACCAACCATCATGG 7049  
QY 1025 CAAAAATGAGGTTTTCTGCGTCCAAACAGAGAGAGGCGCAAGCCAGCTCGCTTA 1084  
Db 7050 CAAAAATGAGGTTTTCTGCGTCCAAACAGAGAGAGGCGCGCAAGCCAGCTCGCTTA 7109  
QY 1085 TCGTGTCCAGACTTGGGGGTCCGTGTGTGCGAAGAAATGACCCTTATGACGTGTCT 1144  
Db 7110 TCGTATTCACAGATTGGGGGTTCGTGTGTGCGAAGAAATGACCCTTATGACGTGTCT 7169  
QY 1145 CCAGCTCCCTCAGGCTGTGATGGGCTCCTCGTACGATTCAGATTTCTCTGACAGC 1204  
Db 7170 CCAGCTCCCTCAGGCGGTGATGGGCTCTTATGAGATTCCAAATCTCTCTGACAGC 7229  
QY 1205 GGGTGAAGTTCCTGTTGAACGCTTGAAATCAAGAGAGCCCTAATGGGCTTGCATATG 1264  
Db 7230 GGGTGAAGTTCCTGTTGAATGCTTGAAAGCGAAGAAATGCCCTATGGGCTTGCATATG 7289  
QY 1265 ACACCCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTAGAAGAGTCAATT 1324  
Db 7290 ACACCCGCTGTTTGAATCAACAGTCACTGAGATGACATCCGTGTAGAAGAGTCAATT 7349  
QY 1325 ATCAATGTTGTACCTTGGGCGCCGGAAGCCAGACAGGCCATAGGTGCTCAAGAGCGGC 1384  
Db 7350 ACCAATGTTGTACTTGGGCGCCGGAAGCCAGACAGGCCATAGGTGCTCAAGAGCGGC 7409  
QY 1385 TTTATATCGGGGGTCCCTGTAATAATTCAAAAGGCGAAGCTGCGGCTATCGCCGATGCC 1444  
Db 7410 TTTATATCGGGGGTCCCTGTAATAATTCAAAAGGCGAAGCTGCGGCTATCGCCGATGCC 7469  
QY 1445 GCGCAGCGGCGTGTGACGACTAGCTGCGTAAATACCTTCAATGTTACTTGAAGGCTT 1504  
Db 7470 GCGCAGCGGCGTGTGACGACTAGCTGCGTAAATACCTTCAATGTTACTTGAAGGCTT 7529  
QY 1505 CTGCAAGCTGTGAGCTGCAAGCTCCAGGACTGCAAGATGCTGTGTGCGGAGACGGCC 1564  
Db 7530 CTGCAAGCTGTGAGCTGCAAGCTCCAGGACTGCAAGATGCTGTGTGCGGAGACGGCC 7589  
QY 1565 TTGTGCTTATCTGTGAGACGCGGGAACCCAGAGAGAGCGGCGAGCTTACGAGTCTTCA 1624  
Db 7590 TTGTGCTTATCTGTGAGAGCGGGAACCCAGAGAGAGCGGCGAGCTTACGAGGCTTCA 7649  
QY 1625 CCGAGGCTATGACTAGTACTTGTGCCCCCGGGGACCGGCCCAACCCAGAAATAGACC 1684  
Db 7650 CCGAGGCTATGACTAGTACTTGTGCCCCCGGGGACCGGCCCAACCCAGAAATAGACC 7709  
QY 1685 TGGAGTTGATACATCATGCTCTCTCAATGTGTGCTGCGGACGATGCTTGGCAAAA 1744  
Db 7710 TGGAGTTGATACATCATGCTCTCTCAATGTGTGCTGCGGACGATGCTTGGCAAAA 7769  
QY 1745 GGGTATACCTCAACCGGTGACC 1769  
Db 7770 GGGTATACCTCAACCGGTGACC 7794



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 144964**

**TO: Bao-Qun Li**  
**Location: rem/3d24/3c18**  
**Art Unit: 1648**  
**Tuesday, February 22, 2005**

**Case Serial Number: 09/664363**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**REM-1A55**  
**Phone: 571-272-2512**

**edward.hart@uspto.gov**

### **Search Notes**

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart



**This Page Blank (uspto)**

3

STIC-Biotech/ChemLib

144964

my

From: Li, Bao-Qun  
Sent: Friday, February 11, 2005 8:12 AM  
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF BASES 308-2116 OF SEQ ID NO; 21 IN APPLICATION SN. 09,664,363. THANKS.

Bao Qun Li M.D  
TC 1600  
Art Unit 1648  
Tel. 517-272-0904  
REM, 3C18  
Rm. 3D24

RECEIVED  
FEB 11 2005  
STIC

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: 2/16/05  
Date Completed: 2/16/05  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA Sequence: # \_\_\_\_\_  
AA Sequence: # \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: 256  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**



# STIC SEARCH RESULTS

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

*Types of relevant prior art found:*

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

**Comments:**

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 21:00:40 ; Search time 8042.09 Seconds  
(without alignments)  
10899.589 Million cell updates/sec

Title: US-09-664-363-21\_COPY\_308\_2116

Perfect score: 1809  
Sequence: 1 ATGAGCAGCAATCTTAAC.....GGGTTGGGGCCTTGTTG 1809

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sta:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1809	100.0	2116	6 A28155	A28155 PT-NANBH MR
2	1809	100.0	2116	6 A32202	A32202 NANBH PT P
3	1809	100.0	2116	6 AR144050	AR144050 Sequence
4	1605.8	88.8	9410	14 HPCX182	D50481 Hepatitis C
5	1593.6	88.1	9359	14 AF313916	AF313916 Hepatitis C
6	1593.6	88.1	9410	14 HPCX182	D50485 Hepatitis C
7	1581.8	87.4	9431	14 HCU45476	U45476 Hepatitis C
8	1581.8	87.4	9585	14 AB049095	AB049095 Hepatitis C
9	1580.2	87.4	9379	14 AF207761	AF207761 Hepatitis C
10	1580.2	87.4	9379	14 AF207766	AF207766 Hepatitis C
11	1575.4	87.1	9547	14 AB049091	AB049091 Hepatitis C
12	1572.2	86.9	9605	6 AX739971	AX739971 Sequence
13	1572.2	86.9	9605	14 HCU238799	HCU238799 Hepatitis C
14	1570.6	86.8	3286	14 AB008441	AB008441 Hepatitis C
15	1570.6	86.8	9033	14 HCU238800	HCU238800 Hepatitis C
16	1570.6	86.8	11076	6 AR406040	AR406040 Sequence
17	1570.6	86.8	11076	6 AR406046	AR406046 Sequence
18	1570.6	86.8	11076	6 AR406048	AR406048 Sequence
19	1570.6	86.8	11076	6 AR406050	AR406050 Sequence

20	1570.6	86.8	11076	6 AX036252	AX036252 Sequence
21	1570.6	86.8	11076	6 AX036258	AX036258 Sequence
22	1570.6	86.8	11076	6 AX036260	AX036260 Sequence
23	1570.6	86.8	11076	6 AX036262	AX036262 Sequence
24	1569.8	86.7	9580	14 AF054250	AF054250 Hepatitis C
25	1568.8	86.7	9410	14 HPCX181	D50480 Hepatitis C
26	1567.4	86.6	9377	14 AF207756	AF207756 Hepatitis C
27	1567.4	86.6	9460	14 HPCX181	D10750 Hepatitis C
28	1567.4	86.6	9533	14 HCU232997	HCU232997 Hepatitis C
29	1565.8	86.6	2340	6 E04260	E04260 cDNA
30	1565.8	86.6	2340	6 E04805	E04805 cDNA to 5'
31	1565.8	86.6	2540	6 E07391	E07391 cDNA
32	1565.8	86.6	9448	14 HPCX183	D13558 Hepatitis C
33	1564.2	86.5	3296	14 AB008442	AB008442 Hepatitis C
34	1564.2	86.5	9456	14 HPCX181	D10934 Hepatitis C
35	1564.2	86.5	9598	14 AB049101	AB049101 Hepatitis C
36	1563.2	86.4	9379	14 AF165052	AF165052 Hepatitis C
37	1562.6	86.4	3296	14 AB008443	AB008443 Hepatitis C
38	1562.6	86.4	8781	14 AF054254	AF054254 Hepatitis C
39	1562.6	86.4	9375	14 AF207754	AF207754 Hepatitis C
40	1562.6	86.4	9379	14 AF165051	AF165051 Hepatitis C
41	1562.6	86.4	9596	14 AF054249	AF054249 Hepatitis C
42	1561	86.3	9573	14 AB049090	AB049090 Hepatitis C
43	1560	86.2	9410	14 HPCX181	D50483 Hepatitis C
44	1559.4	86.2	8780	14 AF054259	AF054259 Hepatitis C
45	1559.4	86.2	9369	14 AF165053	AF165053 Hepatitis C

#### ALIGNMENTS

RESULT 1	A28155	2116 bp	DNA	linear	PAT 07-JUN-1995
LOCUS	A28155	PT-NANBH mRNA fragment from patent GB239245.			
DEFINITION	PT-NANBH mRNA fragment from patent GB239245.				
ACCESSION	A28155				
VERSION	A28155.1	GI:1248638			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
JOURNAL					
FEATURES					
source					
CDS					
ORIGIN					
Query Match	100.0%	Score 1809;	DB 6;	Length 2116;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 1809;	Conservative	0;	Mismatches	0;	Gaps 0;

QY 1 ATGAGCAGAAATCTTAAACCTCAAGAAAAACCAACGTAACCAACCGCCGCCACAG 60  
Db 308 ATGAGACGAATCTTAAACCTCAAGAAAAACCAACGTAACCAACCGCCGCCACAG 367  
QY 61 GAGCTCAAGTTCCCGGCGGTGTGATGATGCTTGTGGAGTTTACTGTTGCCGCGCAG 120  
Db 368 GAGCTCAAGTTCCCGGCGGTGTGATGATGCTTGTGGAGTTTACTGTTGCCGCGCAG 427  
QY 121 GAGCCAGGTTGGGTGTGGCGCGATAGGAAGACTTCGAGCGGTCCGAACCTGTGTA 180  
Db 428 GAGCCAGGTTGGGTGTGGCGCGATAGGAAGACTTCGAGCGGTCCGAACCTGTGTA 487  
QY 181 AGGCGCAACACTATCCCAAGGCTCGCAGCCGAGGCGAGGAGCTGAGCCGCG 240  
Db 488 AGGCGCAACACTATCCCAAGGCTCGCAGCCGAGGCGAGGAGCTGAGCCGCG 547  
QY 241 TACCTTGGCCCTCTATAGGCAACGAGGCGATGAGGTTGGGAGAGTGTCTCTGTCACC 300  
Db 548 TACCTTGGCCCTCTATAGGCAACGAGGCGATGAGGTTGGGAGAGTGTCTCTGTCACC 607  
QY 301 CGTGGCTCCGGCTAGTGGGGCCCACTGACCCCGCGGTAGGTGCGGTAAATTGGGT 360  
Db 608 CGTGGCTCCGGCTAGTGGGGCCCACTGACCCCGCGGTAGGTGCGGTAAATTGGGT 667  
QY 361 AAAGTCATGATACCTCAATGCGGCTTGCGCGACTTCATGAGGTAATTCCGCTCGTC 420  
Db 668 AAAGTCATGATACCTCAATGCGGCTTGCGCGACTTCATGAGGTAATTCCGCTCGTC 727  
QY 421 GGGCTCCCTTAGGGGGCGCTGCGAGGCGCTGCGCGATGCGGTCCGGGTTCTGAGAGAC 480  
Db 728 GGGCTCCCTTAGGGGGCGCTGCGAGGCGCTGCGCGATGCGGTCCGGGTTCTGAGAGAC 787  
QY 481 GGGCGAATAGCAACAGGGAATTTACCGGGTGTCTTCTCTATCTTCTCTGCTGGT 540  
Db 788 GGGCGAATAGCAACAGGGAATTTACCGGGTGTCTTCTCTATCTTCTCTGCTGGT 847  
QY 541 TTGCTGCTCTGTTTACCAATTCAGCTTCGCTTATGAAATGCGCAACGTCGCCGATC 600  
Db 848 TTGCTGCTCTGTTTACCAATTCAGCTTCGCTTATGAAATGCGCAACGTCGCCGATC 907  
QY 601 TACATGTACGAACGATGCTTCAATCTAGAGATGCTGTACGAGACAGCGGACATGATC 660  
Db 908 TACATGTACGAACGATGCTTCAATCTAGAGATGCTGTACGAGACAGCGGACATGATC 967  
QY 661 ATGCAACCCCGGGGTGTGCGCTGTGTCGCGGAGGGAATTCCTCCCGCTGCGGTA 720  
Db 968 ATGCAACCCCGGGGTGTGCGCTGTGTCGCGGAGGGAATTCCTCCCGCTGCGGTA 1027  
QY 721 GCGCTCACTCCACGCTCGCGGCAAGAGCGCAGAGATCCCACTGCGACATACGACGC 780  
Db 1028 GCGCTCACTCCACGCTCGCGGCAAGAGCGCAGAGATCCCACTGCGACATACGACGC 1087  
QY 781 CAGCTGATTTGCTGTGGGGGCTGCTTCTGCTCGCTATGTACGTGGGGATCTC 840  
Db 1088 CAGCTGATTTGCTGTGGGGGCTGCTTCTGCTCGCTATGTACGTGGGGATCTC 1147  
QY 841 TGGGATCTGTTTCTGCTCTCTAGCTGTTCACCTTCCGCTCGCGAATACGAGC 900  
Db 1148 TGGGATCTGTTTCTGCTCTCTAGCTGTTCACCTTCCGCTCGCGAATACGAGC 1207  
QY 901 GTACAGAGATGCAATTTCAATCTATCCCGGCAAGTACAGTCAACGATGCTTGG 960  
Db 1208 GTACAGAGATGCAATTTCAATCTATCCCGGCAAGTACAGTCAACGATGCTTGG 1267  
QY 961 GATATGATGATGAATCTGTACCTTACGACAGCCCTTAGTGTATCGAGCTACTCCGATC 1020  
Db 1268 GATATGATGATGAATCTGTACCTTACGACAGCCCTTAGTGTATCGAGCTACTCCGATC 1327  
QY 1021 CCAAGAGCTGTGTGACATGTGTGGGGGGCCCACTGGGAGATCTGTGGCGGCTTGGC 1080  
Db 1328 CCAAGAGCTGTGTGACATGTGTGGGGGGCCCACTGGGAGATCTGTGGCGGCTTGGC 1387

QY 1081 TACTATTCATGATGGGGAATGAGGCTTGTGTTGTGATGCTACTCTTTGCCGCGC 1140  
Db 1388 TACTATTCATGATGGGGAATGAGGCTTGTGTTGTGATGCTACTCTTTGCCGCGC 1447  
QY 1141 GTTGAAGGGGAACCTTACACGACAGGGGGGACACACGGCCGCGCCACAGGGCTTACA 1200  
Db 1448 GTTGAAGGGGAACCTTACACGACAGGGGGGACACACGGCCGCGCCACAGGGCTTACA 1507  
QY 1201 TCCCTCTTACACACCTGGGCGGCTCAGAAAATTCAGAGTTTGTAAACCAACGGGAGCTGG 1260  
Db 1508 TCCCTCTTACACACCTGGGCGGCTCAGAAAATTCAGAGTTTGTAAACCAACGGGAGCTGG 1567  
QY 1261 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCAAACTGGGTTCTTGGCGCG 1320  
Db 1568 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCAAACTGGGTTCTTGGCGCG 1627  
QY 1321 CTGTTCTACACGACAGGTTCAATGCGTCCGGAATGCTCAGAGCGCATGGCCAGCTGCC 1380  
Db 1628 CTGTTCTACACGACAGGTTCAATGCGTCCGGAATGCTCAGAGCGCATGGCCAGCTGCC 1687  
QY 1381 CCATGTACGATTCGATCAGGGGGGTGCCATCACTTAAATGAGTCCACGCGCTTG 1440  
Db 1688 CCATGTACGATTCGATCAGGGGGGTGCCATCACTTAAATGAGTCCACGCGCTTG 1747  
QY 1441 GACCAAGGCGCTTATTTGCTGCACTACGACCTCAACCGGTGTGATGTCGCCGCTTG 1500  
Db 1748 GACCAAGGCGCTTATTTGCTGCACTACGACCTCAACCGGTGTGATGTCGCCGCTTG 1807  
QY 1501 CAGGTGTGTGGCCCAATGTACTGTTTCACTCAGAGCCCTGTTGTGTGTGGGACGACGAT 1560  
Db 1808 CAGGTGTGTGGCCCAATGTACTGTTTCACTCAGAGCCCTGTTGTGTGTGGGACGACGAT 1867  
QY 1561 CGTTTGGCGGCGCTTACGATCAGATGGGGGTGAGATGAGACGAGCGTGTCTCTCAAC 1620  
Db 1868 CGTTTGGCGGCGCTTACGATCAGATGGGGGTGAGATGAGACGAGCGTGTCTCTCAAC 1927  
QY 1621 AACACGCGCGCCGCAACGCGGGCAACTGTTCGCTGTACATGATGATGATGACCGGGTTC 1680  
Db 1928 AACACGCGCGCCGCAACGCGGGCAACTGTTCGCTGTACATGATGATGATGACCGGGTTC 1987  
QY 1681 ACCAAGAGTGTGGGGGCGCCCGCTGCAACATCGGGGGGTTCGGCACAAACTTTGATC 1740  
Db 1988 ACCAAGAGTGTGGGGGCGCCCGCTGCAACATCGGGGGGTTCGGCACAAACTTTGATC 2047  
QY 1741 TGCCCAAGGACTGTTCCGGAAGATCCGAGGCACTTACACCAATGCGGTTGGGG 1800  
Db 2048 TGCCCAAGGACTGTTCCGGAAGATCCGAGGCACTTACACCAATGCGGTTGGGG 2107  
QY 1801 CTTGGTTG 1809  
Db 2108 CTTGGTTG 2116

RESULT 2  
A32202 2116 bp DNA linear PAT 10-DEC-1996  
LOCUS NANBHV PT polypeptide fragment.  
DEFINITION A32202  
ACCESSION A32202 GI:1926562  
VERSION A32202.1  
KEYWORDS  
SOURCE  
ORGANISM Non-A, non-B hepatitis virus  
Non-A, non-B hepatitis virus  
Viruses; unclassified viruses.  
REFERENCE  
1 (bases 1 to 2116)  
AUTHORS  
JOURNAL Patent: FR 2655990-A 21 21-JUN-1991;  
FEATURES  
location/Qualifiers  
1..2116  
/organism="Non-A, non-B hepatitis virus"  
/db\_xref="taxon:12440"  
308..>2116  
/codon\_start=1  
CDS



/product="partial\_NANBH-Pr polypeptide"  
/protein\_id="CAA02094.1"  
/db\_xref="GI:1926563"  
/translation="MSTNPKPQKTKRNTNRPRPDVKFPGGQIVGVYLLPRGRRL  
GVYATKTSERQPRGRQPIPRARQPERAWAOPYMPVLGNEMGMAVLLSPRG  
SPSPGPTDPRRRSRRLKVIDTLTCGFLDMGYIPLVAPALGABALAHAGVTH  
GVNVAITGNLPGSPSIFILALISCLTIPASAYEVNRSIYHVNDSNSSIYVEDAD  
MIMHPGCVPCYREGNSRCRWALPTLTAAPASITPATIRRHVDLLVGAAPCSMY  
VDDCGSVFLVSQLFPSRHHQIVDDCSITPGHSGHMRMDMMNPSYALGV  
SQLRIPOAVVDMVAHAWGLAVGLAYVMGNMAVLLVMLLPAVDEPPTLGGTH  
GHAAGLISLFTPGPAOKIQLVYTNQSMHINRTALNDSLGTGFLAALFYTHFNAS  
GSEBMAACRPIDPOFGWGPITVYNSHLDORPCMHYAPDQGVLPALQVCPYVC  
FSPVYVGTTRFQAPTRKMGENTDVLINTRBPBRNMWGCITMNSTGTFTKCGG  
PCNIGGVGNNTLICPTDFRHPATYTKCSGSPWL"

## ORIGIN

Query Match 100.0%; Score 1809; DB 6; Length 2116;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 ATGAGCAGAACTCTTAACCTTAAGAAAAACCAAGCTAACCAACCGGCCCAACG 60  
Db ATGAGCAGAACTCTTAACCTTAAGAAAAACCAAGCTAACCAACCGGCCCAACG 367  
Qy 61 GAAGTCAATTCCTCCGGCGGTGTGATCGTTGGTGGAGTTTACTGTTGGCGGACG 120  
Db 368 GAGCTAAGTTCCTCCGGCGGTGTGATCGTTGGTGGAGTTTACTGTTGGCGGACG 427  
Qy 121 GAGCCAGAGTTGGGTGTGCGCGCACTAAGAGACTTCGAGCGGTGCAACTCGTGA 180  
Db 428 GAGCCAGAGTTGGGTGTGCGCGCACTAAGAGACTTCGAGCGGTGCAACTCGTGA 487  
Qy 181 AGGCGCAACCTATCCCAAGGCTGCCAGCCGAGGCGAGGCGCTGGCTCAGCCGG 240  
Db 488 AGGCGCAACCTATCCCAAGGCTGCCAGCCGAGGCGAGGCGCTGGCTCAGCCGG 547  
Qy 241 TACCTTGGCCCTCTATAGGCAAGAGGCGATGGGGTGGGAGAGATGGCTCTGTACCC 300  
Db 548 TACCTTGGCCCTCTATAGGCAAGAGGCGATGGGGTGGGAGAGATGGCTCTGTACCC 607  
Qy 301 CGTGGCTCCGGCTAGTGGGCGCCCACTGACCCCGGCGTATGGCTATTTGGGT 360  
Db 608 CGTGGCTCCGGCTAGTGGGCGCCCACTGACCCCGGCGTATGGCTATTTGGGT 667  
Qy 361 AAAGTCATGATACCTCTCAATGCGGCTTCGCGAAGCTTCAATGGGATTAATTCGCTGTC 420  
Db 668 AAAGTCATGATACCTCTCAATGCGGCTTCGCGAAGCTTCAATGGGATTAATTCGCTGTC 727  
Qy 421 GGGCGTCTCTTAGGGGGCGCTGCCAGGGCGCTGGCGCATGGCGTCCGGGTTCTGAGAGAC 480  
Db 728 GGGCGTCTCTTAGGGGGCGCTGCCAGGGCGCTGGCGCATGGCGTCCGGGTTCTGAGAGAC 787  
Qy 481 GGGCGTGAATATGCAACAGGGAATTAACCGGTTGCTTTCTATCTTCTCTTGGGT 540  
Db 788 GGGCGTGAATATGCAACAGGGAATTAACCGGTTGCTTTCTATCTTCTCTTGGGT 847  
Qy 541 TTGCTGTCTGTGTAACAATTCAGCTTCGCTTATGAATGCGCAAGTGTCCGGATC 600  
Db 848 TTGCTGTCTGTGTAACAATTCAGCTTCGCTTATGAATGCGCAAGTGTCCGGATC 907  
Qy 601 TACATGTCAAGAACATTTGCTCAACTCAAGATGTGTACAGACAGCGGACATGATC 660  
Db 908 TACATGTCAAGAACATTTGCTCAACTCAAGATGTGTACAGACAGCGGACATGATC 967  
Qy 661 ATGCAACACCCGGGGGTGTGCGCTGTGTCGGGAGGGATTAATTCCTCCGCTGCGGGTA 720  
Db 968 ATGCAACACCCGGGGGTGTGCGCTGTGTCGGGAGGGATTAATTCCTCCGCTGCGGGTA 1027  
Qy 721 GGGCTCACTCCAGCGCTCGCGGCAAGAGCGCCAGATCCCACTGCGCAATATACAGCG 780  
Db 1028 GGGCTCACTCCAGCGCTCGCGGCAAGAGCGCCAGATCCCACTGCGCAATATACAGCG 1087  
Qy 781 CAGGTGATTTGCTCGTTGGGGCGGCTGCTTCTGCTCCGCTATGTACGTGGGGGATCTC 840

Db 1088 CAGGTGATTTGCTCGTTGGGGCGGCTGCTTCTGCTCCGCTATGTACGTGGGGGATCTC 1147  
Qy 841 TGGGATCTGTTTCTCTGCTCTCAAGCTGTTCACTTCTCGCTCGCGAATCAAGCG 900  
Db 1148 TGGGATCTGTTTCTCTGCTCTCAAGCTGTTCACTTCTCGCTCGCGAATCAAGCG 1207  
Qy 901 GTACAGATCTGAATTTGTTCAATCTATCCCGGCGAGTACAGGTACCGGATGCTTGG 960  
Db 1208 GTACAGATCTGAATTTGTTCAATCTATCCCGGCGAGTACAGGTACCGGATGCTTGG 1267  
Qy 961 GATATGATGATGAACCTGTCACCTCAAGAGGCGCTAGTGTATCGGAGTACTCCGGATC 1020  
Db 1268 GATATGATGATGAACCTGTCACCTCAAGAGGCGCTAGTGTATCGGAGTACTCCGGATC 1327  
Qy 1021 CCACAAAGCTGTGTGTGACATGTGTGGCGGGGCGCACTGGGAGTCTCGCGGCTTGGC 1080  
Db 1328 CCACAAAGCTGTGTGTGACATGTGTGGCGGGGCGCACTGGGAGTCTCGCGGCTTGGC 1387  
Qy 1081 TACTATTCATGTGTGGGAACTGGGCTAAGTCTTGTGTGTATGCTACTCTTGGCGGC 1140  
Db 1388 TACTATTCATGTGTGGGAACTGGGCTAAGTCTTGTGTGTATGCTACTCTTGGCGGC 1447  
Qy 1141 GTTGAAGGGAACTTACACAGACAGGGGGGACACAGCGCGCGCCACGAGGCTTACA 1200  
Db 1448 GTTGAAGGGAACTTACACAGACAGGGGGGACACAGCGCGCGCCACGAGGCTTACA 1507  
Qy 1201 TCCCTCTTCAACCTGGGCGGCTCGAAGAAATCAAGCTTGTAAACCAACGCGACGTGG 1260  
Db 1508 TCCCTCTTCAACCTGGGCGGCTCGAAGAAATCAAGCTTGTAAACCAACGCGACGTGG 1567  
Qy 1261 CACATTAACAGAACTGCTTGAATCTGAATGACTCCTCTCAAACTGGGTTCTTGGCGGG 1320  
Db 1568 CACATTAACAGAACTGCTTGAATCTGAATGACTCCTCTCAAACTGGGTTCTTGGCGGG 1627  
Qy 1321 CTGTTCTACAGCAAGGTTCAATGCGTCCGAGTGTCTGAGCGGCACTGCGACGCTGCGC 1380  
Db 1628 CTGTTCTACAGCAAGGTTCAATGCGTCCGAGTGTCTGAGCGGCACTGCGACGCTGCGC 1687  
Qy 1381 CCCATTGACAGATTGATCAGGGGATGGGGTCCCATCTATATATGATCCCAAGCTTGG 1440  
Db 1688 CCCATTGACAGATTGATCAGGGGATGGGGTCCCATCTATATATGATCCCAAGCTTGG 1747  
Qy 1441 GACCAAGGCGCTTATGCTGCGCATACGCACTCAACCGTGTGTATGCTGCCCGCTTGG 1500  
Db 1748 GACCAAGGCGCTTATGCTGCGCATACGCACTCAACCGTGTGTATGCTGCCCGCTTGG 1807  
Qy 1501 CAGGTGTGTGGCCAGTGTACGTTTCACTCCAAGCCGTTGTGGTGGGGACAGCGAT 1560  
Db 1808 CAGGTGTGTGGCCAGTGTACGTTTCACTCCAAGCCGTTGTGGTGGGGACAGCGAT 1867  
Qy 1561 CGTTTCGGCGCCCTTACGTAAGATGGGATGAAATGAGACGAGCTGTGCTTCTCAAC 1620  
Db 1868 CGTTTCGGCGCCCTTACGTAAGATGGGATGAAATGAGACGAGCTGTGCTTCTCAAC 1927  
Qy 1621 AACAGCGGCGCCCAAGGGGCAACTGTGTGCGCTGTACATGATGAAATGCAACCGGCTTC 1680  
Db 1928 AACAGCGGCGCCCAAGGGGCAACTGTGTGCGCTGTACATGATGAAATGCAACCGGCTTC 1987  
Qy 1681 ACCAAGACGTTGGGGGCGCCCGCGGCAACATCGGGGGGTGGGCAACCACTTTGATC 1740  
Db 1988 ACCAAGACGTTGGGGGCGCCCGCGGCAACATCGGGGGGTGGGCAACCACTTTGATC 2047  
Qy 1741 TGCCCAAGGAGTGTCTTCGGAAGATCCCGAGGCACTTACACCAATGCGGTTGCGGG 1800  
Db 2048 TGCCCAAGGAGTGTCTTCGGAAGATCCCGAGGCACTTACACCAATGCGGTTGCGGG 2107  
Qy 1801 CCTTGGTTG 1809  
Db 2108 CCTTGGTTG 2116

RESULT 3

ARI44050  
LOCUS ARI44050 2116 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 21 from patent US 6210675.  
ACCESSION ARI44050  
VERSION ARI44050.1 GI:15105917  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2116)  
AUTHORS Highfield,P.Edmund., Rodgers,B.Colin., Tedder,R.Seton. and  
Barbara,J.Anthony.James.  
TITLE PT-NANB hepatitis polypeptides  
JOURNAL Patent: US 6210675-A 21 03-APR-2001;  
FEATURES  
source 1.. 2116  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 100.0%; Score 1809; DB 6; Length 2116;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGCAGCAATCTTAAACCTCAAGAAAACCAACGTAACCAACCGCCGCCACAG 60  
DB 308 ATGAGCAGCAATCTTAAACCTCAAGAAAACCAACGTAACCAACCGCCGCCACAG 367  
QY 61 GACGTCAGATTCCCGGCGGCTGTGATCGATTGGTGAATTACTCTGTCCCGGCAAG 120  
DB 368 GACGTCAGATTCCCGGCGGCTGTGATCGATTGGTGAATTACTCTGTCCCGGCAAG 427  
QY 121 GAGCCAGGTTGGGTGTGCGCGCACTAGAGAAGACTTCGACGCGGTGCAACTGTGGA 180  
DB 428 GAGCCAGGTTGGGTGTGCGCGCACTAGAGAAGACTTCGACGCGGTGCAACTGTGGA 487  
QY 181 AGGCGACAACCTATCCCAAGGCTGCGCAGCCGAGGGGCAAGGCTCTGAGCCCGAG 240  
DB 488 AGGCGACAACCTATCCCAAGGCTGCGCAGCCGAGGGGCAAGGCTCTGAGCCCGAG 547  
QY 241 TACCCCTGGCCCTCTATGAGCAACGAGGCAATGGGTGAGAGATGGTCTCTGTCAACC 300  
DB 548 TACCCCTGGCCCTCTATGAGCAACGAGGCAATGGGTGAGAGATGGTCTCTGTCAACC 607  
QY 301 CGTGGCTCCGGGCTAGTGGGGCCCACTGACCCCGGCGTAGGTCGCGTAATTGGGT 360  
DB 608 CGTGGCTCCGGGCTAGTGGGGCCCACTGACCCCGGCGTAGGTCGCGTAATTGGGT 667  
QY 361 AAAGTCATGATACCTCTCATGCGGCTTGCGCAACCTTCATGGGGTACATTCGCTCGTC 420  
DB 668 AAAGTCATGATACCTCTCATGCGGCTTGCGCAACCTTCATGGGGTACATTCGCTCGTC 727  
QY 421 GGGCGTCCCTTATGAGGGGCGCTGCCAGGGCCCTGGGGCATGGGTCGCGGTTCTGAGAGAC 480  
DB 728 GGGCGTCCCTTATGAGGGGCGCTGCCAGGGCCCTGGGGCATGGGTCGCGGTTCTGAGAGAC 787  
QY 481 GGCCTGAACATATGCAACAGGGAATTTACCCGGTTGCTCTTCTCATCTCTCTGGGT 540  
DB 788 GGCCTGAACATATGCAACAGGGAATTTACCCGGTTGCTCTTCTCATCTCTCTGGGT 847  
QY 541 TTGCTGCTCTGTTTATGCAATTCAGCTTCGCTTATGAAATGCGGCAAGTGTCCGGAGTC 600  
DB 848 TTGCTGCTCTGTTTATGCAATTCAGCTTCGCTTATGAAATGCGGCAAGTGTCCGGAGTC 907  
QY 601 TACCATATGCAAGAAATTTGCTCCAACTAAGCATCTGTATACGAAACAGCGGACATGATC 660  
DB 908 TACCATATGCAAGAAATTTGCTCCAACTAAGCATCTGTATACGAAACAGCGGACATGATC 967  
QY 661 ATGCAACACCCCGGGGTGTGCTGTGTCCGGAGGGGTAATTCCTCCCGGCTGTGGGTA 720  
DB 968 ATGCAACACCCCGGGGT 1027  
QY 721 GCGCTCACTCCACGCTGCGGGCAAGAGCGCAGAGATCCCACTGCGCAATATGAGAGCG 780

DB 1028 GCGCTCACTCCACGCTGCGGGCAAGAGCGCAGAGATCCCACTGCGCAATATGAGAGCG 1087  
QY 781 CACGTCGATTTTGTCTGTTGGGGCGGCTGCTTCTGCTCCGCTATGATAGTGGGGATCTTC 840  
DB 1088 CACGTCGATTTTGTCTGTTGGGGCGGCTGCTTCTGCTCCGCTATGATAGTGGGGATCTTC 1147  
QY 841 TCGGATCTGTTTCTCTGCTCTGCTAGCTGTTCACTCTCGCTCGCGCATATGACAG 900  
DB 1148 TCGGATCTGTTTCTCTGCTCTGCTAGCTGTTCACTCTCGCTCGCGCATATGACAG 1207  
QY 901 GTACAGACTGCAATTTGTTCAATCTATCCGCGCAAGTATCAGGTCAACGCAATGCTTGG 960  
DB 1208 GTACAGACTGCAATTTGTTCAATCTATCCGCGCAAGTATCAGGTCAACGCAATGCTTGG 1267  
QY 961 GATATGATGATGAACCTGTGCACTTACAGACCCCTAGTGTATGCGAGCTACCTCGGATC 1020  
DB 1268 GATATGATGATGAACCTGTGCACTTACAGACCCCTAGTGTATGCGAGCTACCTCGGATC 1327  
QY 1021 CCACAGCTGTCTGTGACATGATGGCGGGGCCCACTGGGAGATCTTGGCGGGCTTGGCC 1080  
DB 1328 CCACAGCTGTCTGTGACATGATGGCGGGGCCCACTGGGAGATCTTGGCGGGCTTGGCC 1387  
QY 1081 TACTATTCATGTGTGGGAACTGGGCTTAAAGTCTTGTGTGATGCTAATCTTTGCCGCG 1140  
DB 1388 TACTATTCATGTGTGGGAACTGGGCTTAAAGTCTTGTGTGATGCTAATCTTTGCCGCG 1447  
QY 1141 GTTGAACGGGAACTTTCACGACAGAGGGGGAACAACGCGCGCGGCCCAAGGGCTTACA 1200  
DB 1448 GTTGAACGGGAACTTTCACGACAGAGGGGGAACAACGCGCGCGGCCCAAGGGCTTACA 1507  
QY 1201 TCCCTCTTCAACCTGGGCGGGCTCAGAAATTCAGCTTGTAAACACCAACGAGAGCTGG 1260  
DB 1508 TCCCTCTTCAACCTGGGCGGGCTCAGAAATTCAGCTTGTAAACACCAACGAGAGCTGG 1567  
QY 1261 CACATCAAACAGACTGCTTGAATCTGCAATGATCTCTCAACCTGGGTTCTTGGCGCG 1320  
DB 1568 CACATCAAACAGACTGCTTGAATCTGCAATGATCTCTCTCAACCTGGGTTCTTGGCGCG 1627  
QY 1321 CTGTTCTACAGCAACAGTTCAATGCGTCCGGAGTCTCAGAGCCCATGCGGACGCGCG 1380  
DB 1628 CTGTTCTACAGCAACAGTTCAATGCGTCCGGAGTCTCAGAGCCCATGCGGACGCGCG 1687  
QY 1381 CCCATTGACAGTTCGATCAAGGGGTGGGGTCCCATCACTTATTAAGTCCCAAGGCTTGG 1440  
DB 1688 CCCATTGACAGTTCGATCAAGGGGTGGGGTCCCATCACTTATTAAGTCCCAAGGCTTGG 1747  
QY 1441 GACCAAGGCCCTTATTTGCTGCACTACGACCTCAACCGTGTGTATCTGTCGGCTTGG 1500  
DB 1748 GACCAAGGCCCTTATTTGCTGCACTACGACCTCAACCGTGTGTATCTGTCGGCTTGG 1807  
QY 1501 CAGGTGTGTGGCCCAAGTACTGTTTCACTCAAGCCCTGTGTGTGTGGGAGCAACCGAT 1560  
DB 1808 CAGGTGTGTGGCCCAAGTACTGTTTCACTCAAGCCCTGTGTGTGTGGGAGCAACCGAT 1867  
QY 1561 CGTTTCGGCGCCCTTACGTAAGATGGGGTGAAGATGAGACGAGCTGCTCTCAAC 1620  
DB 1868 CGTTTCGGCGCCCTTACGTAAGATGGGGTGAAGATGAGACGAGCTGCTCTCAAC 1927  
QY 1621 AACACGGGCGGCAACGGGGCAACTGTTGGCTGTATCATGATGAAATGACACGGGTTTC 1680  
DB 1928 AACACGGGCGGCAACGGGGCAACTGTTGGCTGTATCATGATGAAATGACACGGGTTTC 1987  
QY 1681 ACCAAGAGTGTGGGGGCCCCCGCTGCAACATCGGGGGGGTGGGCAACAACATTTGATC 1740  
DB 1988 ACCAAGAGTGTGGGGGCCCCCGCTGCAACATCGGGGGGGTGGGCAACAACATTTGATC 2047  
QY 1741 TGCCCCACGACTCTTCCGGAACATCCGAGGCACTTACACCAATATGCGGTTGGGG 1800  
DB 2048 TGCCCCACGACTCTTCCGGAACATCCGAGGCACTTACACCAATATGCGGTTGGGG 2107  
QY 1801 CTTGGTTTG 1809



Db 510 AGCGCAACCTATCCCCAAGGCTGCGACGCCGAGGCGAGGCGCTGAGCTCAGCCCGG 569  
Qy 241 TACCTTGGCCCCCTCTATGAGCAACGAGGCGATGGGGTGGGAGAAATGCTCTGTCACCC 300  
Db 570 TACCTTGGCCCCCTCTATGAGCAATGAGGCGATGGGGTGGGAGAAATGCTCTGTCACCC 629  
Qy 301 GGTGGCTCCGGGCTAGTTGGGGCCCGACATGACCCCCCGGCGTAGGTGCGGTAAATTTGGGT 360  
Db 630 GCGGCTCTGGGCTAGTTGGGGCCCGACGAGACCCCCCGGCGTAGGTGCGGTAAATTTGGGT 689  
Qy 361 AAAGTCATGATACCTCTCATGCGGCTTCCCGACCTATAGGGGTACTTCCGCTCGTC 420  
Db 690 AAGGTCATGATACCTCTCATGCGGCTTCCCGACCTATAGGGGTACTTCCGCTCGTC 749  
Qy 421 GCGCTCTCCCTTAGGGGGGCGCTGCAAGGGCCCTGGGCGATGGGTCGCGGTTCTGAGAGAC 480  
Db 750 GCGGCCCCCTTAGGGGGGCGCTTCCAGGGCCCTGGGCGATGGGTCGCGGTTCTGAGAGAC 809  
Qy 481 GCGGCTAACTATGCAACAGGGAATTTAACCGGTTGCTTTCTCTATCTTCCCTTGGCT 540  
Db 810 GCGGCTAACTATGCAACAGGGAATCTGCGCGTGTCTTTCTCTATCTTCCCTTGGCT 869  
Qy 541 TTGCTGTCTGTGTTTGAACCACTTCCAGCTTCCGCTTATGAAGTGGCGAAGTGTCCGAGTC 600  
Db 870 TTGCTGTCTGTGTTTGAACCACTTCCAGCTTCCGCTTATGAAGTGGCGAAGTGTCCGAGTC 929  
Qy 601 TACCATGTCAAGAAAGATTTGCTCAACTCAAGCATCGTGTACAGACAGCGGACATGATC 660  
Db 930 TACCATGTCAAGAAAGATGCTCTTAAACGACATGTTGTATGAGAGCGAGCAATGATC 989  
Qy 661 ATGCAACACCCCGGGGTGTGTGCTGTGTCTCGGAGGAGGTAAATCTCCGCTGTGGGTA 720  
Db 990 ATGCAATACCCCGGGGTGTGTGCTGTGTCTCGGAGGAGCAATCTCCGCTGTGGGTA 1049  
Qy 721 GCGCTCACTCCGAGGCTGGCGGCGCAAGAGCGGACATCCCGACGCGCAATAGAGGC 780  
Db 1050 GCGCTCACTCCGAGGCTGGCGGCGCAAGAGCTCCAGCTCCCAATAGCAATAGAGGC 1109  
Qy 781 CACGTGATTTGCTGTTGGGGGCGCTGCTCTGCTCCGATATGATGGGGATCTC 840  
Db 1110 CACGTGATTTGCTGTTGGGGGCGCTGCTCTGCTCCGATATGATGGGGATCTC 1169  
Qy 841 TGGGATCTGTTTCTCGTCTCTGAGCTGTTACCTTCTGCGCTCGCGCATGACAG 900  
Db 1170 TGGGATCTGTTTCTCGTCTCTGAGCTGTTACCTTCTGCGCTCGCGCATGACAG 1229  
Qy 901 GTACAGGAGCTGAATTTGTTCAATCTATCCCGGCGACATAGGTACCGCATGCTTGG 960  
Db 1230 GTACAGGAGCTGAATTTGTTCAATCTATCCCGGCGACATAGGTACCGCATGCTTGG 1289  
Qy 961 GATATGATGATGAATGTCACCTTACAGAGCCCTAGTGTATCGCAGTACTCCGATC 1020  
Db 1290 GATATGATGATGAATGTCACCTTACAGAGCCCTAGTGTATCGCAGTACTCCGATC 1349  
Qy 1021 CCACAAGCTGTGTGAGAGATGCTGGGGGCGCACTGGGAGTCTCGCGGCGCTTCC 1080  
Db 1350 CCACAAGCTGTGTGAGAGATGCTGGGGGCGCACTGGGAGTCTCGCGGCGCTTCC 1409  
Qy 1081 TACTATTCAATGATGGGAACTGGGCTTAAAGTCTTGGTGTGATGCTACTCTTTGGCGGC 1140  
Db 1410 TACTATTCAATGATGGGAACTGGGCTTAAAGTCTTGGTGTGATGCTACTCTTTGGCGGC 1469  
Qy 1141 GTTGAAGGGGAACCTTACACGACAGGGGGGACACAGGCGCGCGCCACAGGGCTTACA 1200  
Db 1470 GTTGAAGGGGAACCTTACGTCACAGGGGGGGGAAAGACCATACCAACCAAGGGCTCGTG 1529  
Qy 1201 TCCCTTTTCAACCTTGGGCGGCTCAGAAATTCAGCTTTGTAACCAACCAACGAGCTGG 1260  
Db 1530 TCCCTTTTGGACCTGGGGCGGAGAGAAATTCAGCTTTGTAATACCAACGAGCTGG 1589  
Qy 1261 CACATCAACAGAACTGCTTGAATGCTGCAATGATCTCCCTTCAAACTGGGTTCTTGGCGGC 1320  
Db 1590 CACATCAACAGAACTGCTTGAATGCTGCAATGATCTCCCTTCAAACTGGGTTCTTGGCGGC 1649

Qy 1321 CTGTTCTACAGCAGAGTTCAATGCTCCGATGCTCAGAGCGCATGCGGCGCC 1380  
Db 1650 CTGTTCTACAGCAGAGTTCAATGCTCCGATGCTCAGAGCGCATGCGGCGCC 1709  
Qy 1381 CCCATTGACAGTTGATCAGGGGTGGGGTCCCATCACTTAATGATGCCACGGCTTG 1440  
Db 1710 TCCATGACAGGTTGCTCAGGGGTGGGGTCCCATCACTTAAGGTGAGCCGGGAGCTCG 1769  
Qy 1441 GACCAAGAGCCCTATGCTGGCATCAGCACTTCAACCGTGTGTAATGTGCGCGCTTG 1500  
Db 1770 GACCAAGAGCCCTATGCTGGCATCAGCACTTCAACCGTGTGTAATGTGCGCGCTTG 1829  
Qy 1501 CAGGTGTGGCCCACTGATGTTTCACTCAAGCCCTGTTGTGGTGGGAGACCGAT 1560  
Db 1830 GAGGTGTGGTCCAGTATATTGCTTCAACCCCAAGCCCTGTTGTGGTGGGAGACCGAT 1889  
Qy 1561 CGTTTGGCGCCCTTACGTACAGATGGGGTGAATGAGACGACGTCGCTCTCAAC 1620  
Db 1890 CGTTTGGCGCCCTTACGTATAGCTGGGGGAGATGAGACGACGTCGCTCTCAAC 1949  
Qy 1621 AACACGCGGCGCCCAAGGAGCACTGTTGGGCTGTATGATGATGATGACACGGGCTTC 1680  
Db 1950 AACACGCGGCGCCTCAAGGCACTGTTGGGCTGTATGATGATGATGACACGGGCTTC 2009  
Qy 1681 ACCAAGACGTGTGGGGGCGCCCGGCAACATCGGGGGGCTCGGCAACACACTTTGATC 1740  
Db 2010 ACCAAGACGTGTGGGGGCGCCCGGCAACATCGGGGGGCTCGGCAACACACTTTGATC 2069  
Qy 1741 TGGCCACGAGCTGTTCCGGAAGCATCCGAGGCGCACTTACACCAATGGCGTTCCGGGG 1800  
Db 2070 TGGCCACGAGCTGTTCCGGAAGCAACCCGGAAGCATCTACCAATGGCGTTCCGGGG 2129  
Qy 1801 CTTGGTTG 1809  
Db 2130 CTTGGTTG 2138

RESULT 5  
AF313916  
LOCUS AF313916 9359 bp mRNA linear VRL 01-JAN-2002  
DEFINITION Hepatitis C virus polyprotein mRNA, partial cds.  
ACCESSION AF313916  
VERSION AF313916.1 GI:18027684  
KEYWORDS  
SOURCE  
ORGANISM  
Hepatitis C virus  
Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
REFERENCE  
1 (bases 1 to 9359)  
Fanning, L.J., Itakura, J., Nagayama, K. and Enomoto, N.  
Characteristics of Hepatitis C viral genome associated with disease progression in a homogeneous patient population  
Unpublished  
2 (bases 1 to 9359)  
Fanning, L.J., Itakura, J., Nagayama, K. and Enomoto, N.  
Direct Submission  
Submitted (17-OCT-2000) Medicine, National University of Ireland, Cork, Hepatitis C Unit, Clinical Sciences Building, Cork University Hospital, Cork, Ireland  
LOCATION/Qualifiers  
1..9359  
/organism="Hepatitis C virus"  
/mol\_type="mRNA"  
/db\_xref="taxon:11103"  
/note="Isolated from source plasma identified as the single source of the iatrogenic infection of Rhesus negative women in Ireland from May 1977-November 1978; exposure was through HCV 1b contaminated anti-D immunoglobulin"  
1..329  
330..>9359  
5'UTR  
CDS  
/codon\_start=1

```
/product="polypeptide"
/peptide_id="A155821.1"
/db_xref="GI:18027685"
/translation="MSTNPKPQRTKRNTRNRDQVFPQGGQIVGGVYLPRGRRL
GVATRTKTSRQPRGRQPIPKARREGTALMQPPIPLVNEBGMAGMGLSPRG
SRPMPTDPRRRSRMLKVIDLTGCGFALMGVILPVAGPLGAGALAHGVRLED
GVNATGNLPGGFSFILALSLCTIPASILEVRNVSQVHYHTDLSNATSEAD
MIMPRGCVPCURENNSRCVATLPTLAARNSIPTTTIRRVVDLVTGAACSMY
VDDLCGVFLVLSQFTSPRHETVDDCSSTYGHVSGHRMMDMMSTTALV
SOLLRIPQAVDMVAGWGLAGLAYSAGMAGVLYMLFAVDVETETTGAGQ
AHTTRGLATLFSFGSQNIOLINTNGSNINRALTNCNLSLNGFALALFYARFAS
GGERMASCRPIDKFOGMPILAKPLSCGCMHAPQCGIVPSVQGYPC
FPSPVVGTTDRFGVPTSMGENETDVLINTRPRMGECSTMMNGSPKTCGG
PRCDIGVNDPLICTPDCFRKRPBATYTKCGSPMLTRCWDYRYRLMRYCTANP
TIFKRVNYGVVHRLNACNMTGERCDEDRSELSPLSTILEMOVLPSFTLL
PALSTGLIHLHRIVDVOYLXIGSAGVSAIKMEVYLLPLFLAARACACMMMLL
IVQAEALLENLVNLAASVAGTGLISFLVFCAMAYIKGRILPGAAVAFYWGPIIL
LILALPRAYAMDREMAASCGAVFGLALTLSPHYKFLALIMWLQVFIKRAAS
LQWILPLNVRGRDALILMCAIHPBLIPATIKILATLPLAMVAGITRYPYVR
AHGILPACVLRVAGHYVQNALMLALGTYYVDHITPLRDMAGRIDLAIVE
PVSFDMETKITTWGADTAAGDIILGLPVSARKREILIGRMDLSEGGMRLAITY
ATAQOTRGLAGCLVSLTGRDNQVGEVQVSTATQSLATVNCQMTVHGAQIT
TLAQGKPIITOMYTNVDOLVGMILPPGASILPTCGSDILVTRHADVIPIVRBG
DSRGLSPRPVSLKSGSGPLCPGSAVGIKRAVCAAGKAVDPEVPSMETT
MSPVTDNSSPRAVQTFQVAILHAPTCGSKSTKVPALVAGYKVLVLPBVAATL
GGAVMKAGHVDNRTGRTITTCAPITTSYTGKFLADGCGSGAYOIIICDEHS
TSTTTLIGTIVDQATNGARLVATIPQSVTPHNEIEVALNSINIGLPIYK
ALPIETIKGGRHLIFCHSRKCDLAKLSGLGLANAVRAGLDVSLIPSGVVVA
TALMTPTGDPDPSVIDCNTCVQIVDFPFIETTVPODAVSRSORRTGSG
RTGIYRFTVGERPSGMPDSVLCCECYDAGCAYELTPETSRYLAAYNTPLPYCO
DHEFWEGVFTGLTHIDAFHLSOTKQAGNPPYLVAVQACRAQAPPMQMMKC
LIRKPTLAGPTPLIRLGAQNEVILTHPIKTFMACSADILEVATSTWVYGVILA
ALAAVLTITGSVIVGRILLSGKPAIIPREVLYQOFDEMEBATHLPYIEQMOLAE
QPKQKMLGLQTAHQEAAAPVESKMALESFMAKHMNFISGQIYAGLSTLBN
PAIASLMAFASITSLPTQHTLFINILGCVWAQALPSSASAFAGIAGAVASI
GKGVLDIILAGGAVAGALVAFKVSQEMSTEDLNLPLIPLISGLVAVGVCA
ILRBHVGBGAVQMMNRLIAPASRGHNSPTHYVPSNAAARVIOILSLITTOIK
RLHOMINDCSTPCSGSLRDWDMLCTVLTDPKTLQSLKPLRGLVPLSGQRKX
GWRSGDLMQITPCGAQITGHVKNSSMIIVGRTCSNHWGPFPIINATVTEGTRR
ABNSRSLMRVIAEYLEVTOVDFHYVMTDNKVCQVPAPFPEVTEGVRSLR
VAPCKPLREETFOVGLNOYLVSGLCEPEPDVAVATLSMLTDSHITAEAKRL
ARGSPSLASSASOLASPLKATCTTRSDPADLILENLMROMGNIRIPESEN
KYVILDSFPLRAEBDEREVSAPALERSKRPKRAMPIIARDVNPVLEAKMDDY
VPPVNGCELPPTKAPRIPPRKPTVULTESTVSALABLATIKTGSSESAVDSGT
AARPQPSDDGSDVESYSMPLEBPGDPLSDSDSMTKVSSEADVDVCSWS
YVTGALITPCSAEBSKLPINALNSLRLHNNVYATIRFSASQKQKTPRLQYLD
DHYRDLKEMKAKASTYKAKLPEBAACLTLPHSKFGACAKDVRMLSSAVNHI
LSWKDLLEDETETPIDTTIMAKNEVFCVQPEKGRKAPLIVPDLGAVCEMALYD
VSTLPOAVMGSSVYGRVSGORVEFLVNAKKKPKMPADVTRCFDSVTVESDRIY
RESIYOCCLAPARQAIRSLERYITGEPILNSKONGYRCRASGVLTISCGNTL
TYLKLASACRAKLDDCTMLVCGDULVVICESAGTXEDASIRVETEMTRSAEG
DPPKEPYDELITSCSNVSVAHDAGSKVYVLRDPTPLPAAAEATMRHPVMSL
GNIIWAPFLMARMILMTHFSSILAOOLEKALDCOIGACISIEPLDLPOLIOQLH
GSAFELSHSYPEIRINRVAACLRKIGVPLRVHRBARVAKLSISGGRANCGYTL
ENMAVRTXKLTPRIPASQLDLSMFWVAGSGDIYHSLRBRPRMFLCLLLLSVG
GYLPLNR"
```

## ORIGIN

```
Query Match      88.1%; Score 1593.6; DB 14; Length 9359;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
```

QY 1 ATGAGCAGCAATCTTAAACCTCAAGAAAAACCAACGTAACCAACCGCGCCCAACG 60  
DB 330 ATGAGCAGCAATCTTAAACCTCAAGAAAAACCAACGTAACCAACCGCGCCCAACG 389

QY 61 GAGCTCAAGTCCCGGCGGTGTGATCGTTGGTGAAGTTAACGTTGGCGGCGAG 120  
DB 390 GAGCTCAAGTCCCGGCGGTGTGATCGTTGGTGAAGTTAACGTTGGCGGCGAG 449

QY 121 GGGCCCAAGTTGGGTGCGCGGCACTAGAGAGACTTCCGAGCGGTGCAACTCGTGA 180  
DB 450 GGGCCCAAGTTGGGTGCGCGGCACTAGAGAGACTTCCGAGCGGTGCAACTCGTGA 509

```
QY 181 AGGCGCAACCTATCCCAAGGCTCGCCAGCCCGGAGGCGAGGCTCGGCTACGCCGG 240  
DB 510 AGGCGCAACCTATCCCAAGGCTCGCCAGCCCGGAGGCGAGGCTCGGCTACGCCGG 569

QY 241 TACCTTGGCCCCCTATGAGCAACGAGGCAATGAGGCTGAGGATGAGCTCTGACCC 300  
DB 570 TACCTTGGCCCCCTATGAGCAATGAGGCAATGAGGCTGAGGATGAGCTCTGACCC 629



QY 301 CGTGAGTCCCGGCTAGTGGGCGCCCACTGACCCCGGCGTAAAGTCCGCTAATTTGG 360  
DB 630 CGTGAGTCCCGGCTAGTGGGCGCCCACTGACCCCGGCGTAAAGTCCGCTAATTTGG 689



QY 361 AAGTATGATGATACCTTCAATGCGGCTTCCGCGACCTCATGAGGATACATTCGCTG 420  
DB 690 AAGTATGATGATACCTTCAATGCGGCTTCCGCGACCTCATGAGGATACATTCGCTG 749



QY 421 GGGCTCCCTTAGGGGGGCGCTGCGAGAGGCTTGGCGCATGGCGTCCGGGTCTGAGAG 480  
DB 750 GGGCTCCCTTAGGGGGGCGCTGCGAGAGGCTTGGCGCATGGCGTCCGGGTCTGAGAG 809



QY 481 GGGGTAATATGCAACAGGAAATTTACCGGCTCTTCTCTATCTTCTTGGCT 540  
DB 810 GGGGTAATATGCAACAGGAAATTTACCGGCTCTTCTCTATCTTCTTGGCT 869



QY 541 TTGCTGCTCTTGAACCAATTCAGCTTCCGCTTATGAGAGTGGCGCAAGTCCGGATC 600  
DB 870 TTGCTGCTCTTGAACCAATTCAGCTTCCGCTTATGAGAGTGGCGCAAGTCCGGATC 929



QY 601 TACATATGACGAACGATGCTCAACTGACATGATGATGACGAGAGCATGATC 660  
DB 930 TACATATGACGAACGATGCTCAACTGACATGATGATGACGAGAGCATGATC 989



QY 661 ATGCAACCCCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
DB 990 ATGCAACCCCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1049



QY 721 GGGCTACCTCCCAAGCTCGCGGCGCAAGAGCGGATGATGATGATGATGATGATG 780  
DB 1050 GGGCTACCTCCCAAGCTCGCGGCGCAAGAGCGGATGATGATGATGATGATGATG 1109



QY 781 CAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 1110 CAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1169



QY 841 TGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
DB 1170 TGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1229



QY 901 GTRACGAGTGCATTTGCTCAATCTATCCCGGCGATGATGATGATGATGATGATG 960  
DB 1230 GTRACGAGTGCATTTGCTCAATCTATCCCGGCGATGATGATGATGATGATGATG 1289



QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
DB 1290 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349



QY 1021 CCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
DB 1350 CCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1409



QY 1081 TACTATTCATGCTGAGGAACTGAGGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1140  
DB 1410 TACTATTCATGCTGAGGAACTGAGGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1469



QY 1141 GTTGAAGGAGAACTTACAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
DB 1470 GTTGAAGGAGAACTTACAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1529



QY 1201 TCCCTCTTCAACCTGAGGCGGCTGAGAAATCAAGCTTGTAAACCAACGAGGAGTGG 1260  
DB 1530 ACCCTCTTCTGTTGGGCGGCTGAGAAATCAAGCTTGTAAACCAACGAGGAGTGG 1589


```

	Accession	Version	Keywords	Source Organism	Reference Authors	Title	Journal
Oy	1261	CACATCAACAAGAACTGCGTTGAACGTGAACTAGACTCCCTTCAAACCTGGGTTTCCTTGCCGGC		Hepatitis C virus (strain HCV-1b, clone HCV-K1-S2), complete genome sequence.	Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Izumi, N., Marumo, F. and Sato, C.	Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region	
Db	1590	CACATCAACAAGAACTGCGTTGAACGTGAACTAGACTCCCTTCAAACCTGGGTTTCCTTGCCGGC		Hepatitis C virus	J. Clin. Invest. 96 (1), 224-230 (1995)		
Oy	1321	CTGTTCTAACACGCACAGGTTCAATGTCGTCGAGTCTCAGACGCCATGAGCAGCTGCCGC		Hepatitis C virus	Enomoto, N.		
Db	1650	CTGTTCTAACACGCACAGGTTCAATGTCGTCGAGTCTCAGACGCCATGAGCAGCTGCCGC		Hepatitis C virus	Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental		
Oy	1381	CCCATTTGACCAAGTTCTGCTCAGAGGGTGGGTCCCACTTAATAAGATGCCACGGCTTG		Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			
Db	1710	CCCATTTGACCAAGTTCTGCTCAGAGGGTGGGTCCCACTTAATAAGATGCCACGGCTTG		Hepacivirus.			
Oy	1441	GACCAGAGGCCCTATTGCTGGCACTACGACCTCAACCGTGTGATCGTCCGCGCTTG		Hepacivirus.			
Db	1770	GACCAGAGGCCCTATTGCTGGCACTACGACCTCAACCGTGTGATCGTCCGCGCTTG		Hepacivirus.			
Oy	1501	CAGTGTGTGGCCCCAGTGTACTGTTTCACTCCAGCCCTGTTGTGTGGGAGCAAGCGAT		Hepacivirus.			
Db	1830	CAGTGTGTGGCCCCAGTGTACTGTTTCACTCCAGCCCTGTTGTGTGGGAGCAAGCGAT		Hepacivirus.			
Oy	1561	CGTTTGGCGCGCCCCTACGTACGATGGGAGTGAAGAAGAGCGAGCGTCTGCTTCTCAC		Hepacivirus.			
Db	1880	CGTTTGGCGCGCCCCTACGTACGATGGGAGTGAAGAAGAGCGAGCGTCTGCTTCTCAC		Hepacivirus.			
Oy	1621	AACACGCGCGCCGACGAGGAGCAACTGTGTTGCGCTGTACATGATGAATAGAACCGGGTTC		Hepacivirus.			
Db	1950	AACACGCGCGCCGACGAGGAGCAACTGTGTTGCGCTGTACATGATGAATAGAACCGGGTTC		Hepacivirus.			
Oy	1681	ACCAAGACGTGTGGGGGGGGCCCCCGCTGCAACATCGGGGGGGGTTCGGCAACAACATTGATC		Hepacivirus.			
Db	2010	ACCAAGACGTGTGGGGGGGGCCCCCGCTGCAACATCGGGGGGGGTTCGGCAACAACATTGATC		Hepacivirus.			
Oy	1741	TGCCCCAACGAGCTGTTCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTCCGGGG		Hepacivirus.			
Db	2070	TGCCCCAACGAGCTGTTCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTCCGGGG		Hepacivirus.			
Oy	1801	CCTTGCTTG 1809		Hepacivirus.			
Db	2130	CCTTGCTTG 2138		Hepacivirus.			
RESULT 6							
HPCKIS2							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE ORGANISM							
REFERENCE AUTHORS							
TITLE							
JOURNAL MEDLINE PUBMED							
REFERENCE AUTHORS JOURNAL REFERENCE AUTHORS TITLE JOURNAL							
Submitted (08-MAY-1995) Nobuyuki Enomoto, Tokyo Medical and Dental							

**FEATURES**

University, Second Department of Internal Medicine; 1-5-45 Yushima Bunkyo-ku, Tokyo 113, Japan (E-mail:PXND4522@fvcserve.or.jp, Tel:03-3813-6111 (ex.3224), Fax:03-3818-7177)

**Location/Qualifiers**

```
source
1. .9410
   /organism="Hepatitis C virus"
   /mol_type="genomic RNA"
   /strain="HCV-1b"
   /db_xref="taxon:11103"
   /clone="HCV-K1-S2"
   /note="IN-sensitive HCV from patient 2"
330. .9362
CDS
```

mat\_peptide  
mat\_deptide



mat\_peptide 1479..2756  
/product="E2"  
mat\_peptide 2757..3407  
/product="NS2"  
mat\_peptide 3408..5300  
/product="NS3"  
mat\_peptide 5301..5462  
/product="NS4A"  
mat\_peptide 5463..6245  
/product="NS4B"  
mat\_peptide 6246..7586  
/product="NS5A"  
mat\_peptide 7587..9359  
/product="NS5B"  
misc\_feature 1479..1559  
/note="Hypervariable region (HVR)"  
misc\_feature 6954..7073  
/note="Interferon sensitivity determining region (ISDR)"  
ORIGIN

Query Match 88.1%; Score 1593; DB 14; Length 9410;  
Best Local Similarity 92.5%; Pred. No. 0;  
Matches 1674; Conservative 0; Mismatches 135; Indels 0; Gaps 0;  
QY 1 ATGAGCAGCAATCTTAACCTCAAGAAAAACCAACGTAACCAACCGCGCCACAG 60  
DB 330 ATAGCAGCAATCTTAACCTCAAGAAAAACCAACGTAACCAACCGCGCCACAG 389  
QY 61 GAGTCAGATTCCCGGCGGTGTGTCAGATGCTGTGTGAGATTACCTGTTCGCGCAG 120  
DB 390 GAGTCAGATTCCCGGCGGTGTGTCAGATGCTGTGTGAGATTACCTGTTCGCGCAG 449  
QY 121 GGGCCAGATTGGGTGTGTCGCGGCTAGAGAACTTCCGAGGGTTCGAACCTCGTGA 180  
DB 450 GGGCCAGATTGGGTGTGTCGCGGCTAGAGAACTTCCGAGGGTTCGAACCTCGTGA 509  
QY 181 AGGCGCAACCTATCCCGCAAGGCTCGCAGCCGAGGGGCTGGGGCTCAGCCCGG 240  
DB 510 AGGCGCAACCTATCCCGCAAGGCTCGCAGCCGAGGGGCTGGGGCTCAGCCCGG 569  
QY 241 TACCTTGGCCCTCTATGCGCAAGAGGATGGGGTGGCAGAGATGGCTCTGTCAACC 300  
DB 570 TACCTTGGCCCTCTATGCGCAAGAGGATGGGGTGGCAGAGATGGCTCTGTCAACC 629  
QY 301 CGTGGCTCCCGGCTGTATTGGGCCCCCACTGACCCCGCGGTAGTGGCGTAATTGGGT 360  
DB 630 CGTGGCTCCCGGCTGTATTGGGCCCCCACTGACCCCGCGGTAGTGGCGTAATTGGGT 689  
QY 361 AAGTCATGATACCTCTCATGCGGCTTGGGACCTCATGGGGTATCATTCGGCTGTC 420  
DB 690 AAGTCATGATACCTCTCATGCGGCTTGGGACCTCATGGGGTATCATTCGGCTGTC 749  
QY 421 GGGCTCCCTTAAGGGGCGCTGCGAGGCGCTGGCGCATGGCTCGGGTTCGAGAGAC 480  
DB 750 GGGCTCCCTTAAGGGGCGCTGCGAGGCGCTGGCGCATGGCTCGGGTTCGAGAGAC 809  
QY 481 GGGGTAACTATGCAACAGGAAATTTAACCCGTTGCTCTTCTCTATCTTCTCTGGCT 540  
DB 810 GGGGTAACTATGCAACAGGAAATTTAACCCGTTGCTCTTCTCTATCTTCTCTGGCT 869  
QY 541 TTGCTGTCCGTTTGAACATTCAGCTTCGCTTAAGAGTGGCAACGATTCGGGATC 600  
DB 870 TTGCTGTCCGTTTGAACATTCAGCTTCGCTTAAGAGTGGCAACGATTCGGGATC 929  
QY 601 TACCATGTACGAAACGATTGCTCAACTCAAGCATGTGTACGAGACAGCGACATATC 660  
DB 930 TATCATGTACGAAACGATTGCTCAACTCAAGCATGTGTATGAGGACAGCGACATATC 989  
QY 661 ATGCAACACCCCGGGGTGTGCTGTGTTCGCGAGGGTATTCCTTCGGCTGTGGGT 720  
DB 990 ATGCAACACCCCGGGGTGTGCTGTGTTCGCGAGGGTATTCCTTCGGCTGTGGGT 1049  
QY 721 GCGCTCACTCCACGCTCGCGGCGCAAGAGACGACGATCCCACTGGACAAATACGACG 780

DB 1050 GCGCTCACTCCACGCTCGCGGCGCAAGAACTCCAGGCTCCCAACTACGAAATACGACG 1109  
QY 781 CAGTCGATTTGTGCTGTGGGGGCGCTGCTTCTGCTCCGCTATGTAGTGGGGATCTC 840  
DB 1110 CAGTCGATTTGTGCTGTGGGGGCGCTGCTTCTGCTCCGCTATGTAGTGGGGATCTC 1169  
QY 841 TGGGATCTGTTTCTGCTCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
DB 1170 TGGGATCTGTTTCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1229  
QY 901 GTACAGACTGCAATGTTCAATCTATCCCGGCAAGTATCAGGTACCGCATGCTTGG 960  
DB 1230 GTACAGACTGCAATGTTCAATCTATCCCGGCAAGTATCAGGTACCGCATGCTTGG 1289  
QY 961 GATATGATGATGAATGTCCTACCTTACAGAGCCCTAGTGTATTCGAGCTACTCGGATC 1020  
DB 1290 GATATGATGATGAATGTCCTACCTTACAGAGCCCTAGTGTATTCGAGCTACTCGGATC 1349  
QY 1021 CCACAGCTGTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
DB 1350 CCACAGCTGTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1409  
QY 1081 TACTATTCATGATGGGAACTGGGCTAAGTCTTGTGTGTATGCTACTTGTGCGGCG 1140  
DB 1410 TACTATTCATGATGGGAACTGGGCTAAGTCTTGTGTGTATGCTACTTGTGCGGCG 1469  
QY 1141 GTTACGAGGAACTTACACGACAGAGGGGACACAGCGCGCGCCACGAGCTTACA 1200  
DB 1470 GTTACGAGGAACTTACACGACAGAGGGGACACAGCGCGCGCCACGAGCTTACA 1529  
QY 1201 TCCCTCTTACACCTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACGCGAGCTGG 1260  
DB 1530 TCCCTCTTACACCTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACGCGAGCTGG 1589  
QY 1261 CACATCAACAGAACTGCTTGAACCTGCAATGACTCCCTCAAACTGGTTCCTTGGCGCG 1320  
DB 1590 CACATCAACAGAACTGCTTGAACCTGCAATGACTCCCTCAAACTGGTTCCTTGGCGCG 1649  
QY 1321 CTGTTCTACAGCAGAGTTCAATGCTCGAGTGTCTGAGCGCATGGCCAGCTGCGC 1380  
DB 1650 CTGTTCTACAGCAGAGTTCAATGCTCGAGTGTCTGAGCGCATGGCCAGCTGCGC 1709  
QY 1381 CCCATTGACCACTTGTATGAGGGGTGGGTCCCATCATCTTAATATGATCCACGCTTG 1440  
DB 1710 TCCATTCAGAGCTTGTGCTCAGGGGTGGGTCCCATCATCTTAATATGATCCACGCTTG 1769  
QY 1441 GACAGAGGCTCTATGCTGAGCACTACGACCTCAACCGTGTGTATGCTGCGCGCTTG 1500  
DB 1770 GACAGAGGCTCTATGCTGAGCACTACGACCTCAACCGTGTGTATGCTGCGCGCTTG 1829  
QY 1501 CAGGTGTGTGCGCAAGTACTGTTTCACTCAAGCCCTGTGTGTGTGGGAGCAACGAT 1560  
DB 1830 GAGGTGTGTGCGCAAGTACTGTTTCACTCAAGCCCTGTGTGTGTGGGAGCAACGAT 1889  
QY 1561 CGTTTGGGCGCCCTACGTAAGATGAGTGGGAGAAATGAGACGAGTGTGCTCTCAAC 1620  
DB 1890 CGTTTGGGCGCCCTACGTAAGATGAGTGGGAGAAATGAGACGAGTGTGCTCTCAAC 1949  
QY 1621 AACAGCGGCGCGCAAGGGGCAACTGGTTCGAGCTGTAACATGAGATGAGCAACCGGCTTC 1680  
DB 1950 AACAGCGGCGCGCGCAAGGGGCAACTGGTTCGAGCTGTAACATGAGATGAGCAACCGGCTTC 2009  
QY 1681 ACCAAGACGTGTGGGGGCGCCCGGTGCAACATCGGGGGGTGGGCAACAACCTTGTATC 1740  
DB 2010 ACCAAGACGTGTGGGGGCGCCCGGTGTAACATCGGGGGGTGGGCAACAACCTTGTATC 2069  
QY 1741 TGGCCCAAGGACTGCTTCCGGAAGCATCCGAGGCACTTAAACCAATGCGGTTTCGGGG 1800  
DB 2070 TGGCCCAAGGACTGCTTCCGGAAGCATCCGAGGCACTTAAACCAATGCGGTTTCGGGG 2129  
QY 1801 CCTGGTTG 1809





OY	721	GGGCTCACATCCCAAGCTGGGGGCGCAAGACGCAACATCCCACTGGGCAATACGACG	780
Db	1062	GCGCTTACCCCAACGCTCGGGGCGAGAAACGCAACATCCCTACTAGCAACATACGAGCT	1121
OY	781	CACGTGATTTGCTGTTGGGGCGGCTGCTTCTGCTCCGCTATGTACGTGGGGATCTC	840
Db	1122	CACGTGATTTGCTGTTGGGGGCGGCTGCTTCTCTCCGCTATGTACGTGGGGATCTC	1181
OY	841	TGGCGATTTGTTTTTCTGCTCTCTCAAGCTGTTCACTTCTGGGCTCGCGGACATCAGACG	900
Db	1182	TGGCGGTCGTGTTTTCTGCTCTCTCAAGCTGTTCACTTCTGGGCTCGCGGACATCAGACA	1241
OY	901	GTAACAGACTCAATTTGTTCAATCTATCCCGGCGCATATCAGGTCAACGATGGCTTGG	960
Db	1242	GTTACAGACTCAATTTGCTCAATCTATCCCGGCGCATATCAGGTCAACGATGGCTTGG	1301
OY	961	GATATGATGATGAATCTGGTCACTTACAGACAGCCGTAAGTGATATCGACGTACTCCGATC	1020
Db	1302	GATATGATGATGAATCTGGTCCGCCACAGCTAGTGATATCGCATTTACTCCGATC	1361
OY	1021	CCACAGCTGTCGTGGACATAGTGGCGGGGCGCATGGGGAGTCCTGGCGGCTTGGC	1080
Db	1362	CCGCAAGCCGTGTGGACATAGTGGCGAGGCGCATGGGGAGTCCTGGCGGCTTGGCC	1421
OY	1081	TACTATTCATGCTGGGGAGACTGGGCTTAAGGCTTGTTGTGATGCTACTCTTTGCGCGC	1140
Db	1422	TACTATTCATGCTGGGGAGACTGGGCTTAAGGCTTGTTGTGATGCTACTCTTTGCGAGGC	1481
OY	1141	GTTGACGGGGAACTTTACACGACAGGGGGAGACACAGCGCGCGGCCCCACAGGGCTTACA	1200
Db	1482	GTTGACGGGAACTTACACCGGTACAGGGGGGGGTGCAAGCTGCATACGTTACGAATGACA	1541
OY	1201	TCCCTCTTCAACACTGGGGCGGAGCTGAGAAAAATCCAGCTTGTATAACCAACGAGAGCTGG	1260
Db	1542	AGCTCTATACACGGGGGCGGCTCTCGAGAAATCCAGCTCTGTAAACCAACGAGAGCTGG	1601
OY	1261	CACATCAACAGAACTGCTTGAATCTGCATGACTTCCCTCCAACTGGGTTCTTTGCGCGG	1320
Db	1602	CACATCAACAGAGCGGCGCTGAACTGTATGACTCCCTCCAGACGAGGGTTCTTTGCGCGG	1661
OY	1321	CTGTTCTTACAGGCAAGGTTCAATCCGCTCCGAGTGTACAGGCGGATGGCCAGCTGCCG	1380
Db	1662	CTGTTCTTACACACGAGCTTCAACAGCTCCGAGATGCCAGAGCGGATGGCCAGCTGCCG	1721
OY	1381	CCCATTTGACAGTTTGCATCAGGGGTGGGGTCCCATCACTTATTAATGATGCCAGGCTTG	1440
Db	1722	TCCATTCGACAGTTTGCATCAGGGGTGGGGTCCCATCACTTATTCGCTGAGCCAGGAGCTTG	1781
OY	1441	GACCAAGAGGCCCTTATTTGCTGGACCTACGCACTTCAACCTGTGGTATCTGTGCCGCGTTG	1500
Db	1782	GACCAAGAGGCCCTTATTTGCTGGACCTACGCGCGCACCTGTGTGATGTATGCCGCGCTG	1841
OY	1501	CAGGTGTGGGCCCAAGTACTGTTTCACTCCAAAGCCCTGTTGTGGTGGGAGACACGAT	1560
Db	1842	CAGGTGTGTGCTCAGGTATTTGCTTCAACCCCAAGCCCCGTTGTGGTGGGAGACACGAT	1901
OY	1561	CGTTTCGGCGGCCCTTACGTCAGATAGGAGTGAATGACGAGAGCTGCTCTCAAC	1620
Db	1902	CGCTTCGGCGGCTCCCTTACGTTAGCTGGGGGAGAAATGACGAGAGCTGCTCTCAAC	1961
OY	1621	AAACAGCGGGCGGCAACGGGGCAACTGGTTTGGGCTGTACATGAGATGAATAGCACGGGGTTC	1680
Db	1962	AAACAGCGGGCGGCAACGGGGCAACTGGTTTGGCTGTACATGAGATGAATAGCACGGGGTTC	2021
OY	1681	ACCAAGAGCTGTGGGGGCGGCCGCTGCAACAATCGGGGGGCTCGGCAACAACTTTGATC	1740
Db	2022	ACCAAGAGCTGTGGGGGCGGCCCATGTGGGAGATCGGGGGGCTCGGCAACAACTTTGATC	2081
OY	1741	TGCCCCAGGACTGTTCCGGAAGATCCGAGGCACTTACACCAATTCGGGTTCCGGG	1800
Db	2082	TGCCCAAGACTGTTCCGGAAGATCCCGAGGCACTTACACCAATGTGGGTTCCGGG	2141

[illegible]

3'UTR  
/note="upstream of poly-U stretch"  
9344..9381  
3'UTR  
/note="poly(U) length is indistinct  
9467..9585  
/note="3' X tail"

ORIGIN

Query Match 87.4%; Score 1581.8; DB 14; Length 9585;  
Best Local Similarity 92.2%; Pred. No. 0;  
Matches 1667; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 1 ATGAGCAGCAATCTTAACTCTAAGAAACCAACGTAACCAACCGCGCCCAACG 60  
DB 311 ATGAGCAGCAAAATCTTAACTCTAAGAAACCAACGTAACCAACCGCGCCCAACG 370  
QY 61 GACGTCAAGTCCCGGCGGTGTCAGATCGTTGGAGATTACCTGTGCGCGCAGG 120  
DB 371 GACGTCAAGTCCCGGCGGTGTCAGATCGTTGGAGATTACCTGTGCGCGCAGG 430  
QY 121 GGCCTTGGGTTGGTGTGCGCGACTAGAGAGATTTCCGACGCTCGCAACTCGTGA 180  
DB 431 GGCCTTGGGTTGGTGTGCGCGACTAGAGAGATTTCCGACGCTCGCAACTCGTGA 490  
QY 181 AGGCGCAACCTATCCCAAGCTGCGCAAGCCGAGGCGACGCTGCGCTCAAGCCGCG 240  
DB 491 AGGCGCAACCTATCCCAAGCTGCGCGCGCGCGAGGCTGCGCTCAAGCCGCG 550  
QY 241 TACCTTGGGCTCTATAGCAAGAGGCGATGGGGTGGGAGAGATGCTCTGTGACCC 300  
DB 551 TACCTTGGGCTCTATAGCAAGAGGCTGCGGGTGGGAGAGATGCTCTGTGACCC 610  
QY 301 CGTGGCTCCGCGCTAGTGGGCGCCCACTGACCCCGCGCTAGGTCGCTAATTTGGGT 360  
DB 611 CGTGGCTCCGCGCTAGTGGGCGCCCACTGACCCCGCGCTAGGTCGCTAATTTGGGT 670  
QY 361 AAGATCATGATACCTCTACATGCGGCTTGGCGCACTCATGAGGCTATCTCGCTGCT 420  
DB 671 AAGATCATGATACCTCTACATGCGGCTTGGCGCACTCATGAGGCTATCTCGCTGCT 730  
QY 421 GCGGCTCCCTTGGGGGCGCTGCGAGGCGCTGCGGCGATGCGGCTTCTGAGAGAC 480  
DB 731 GCGGCTCCCTTGGGGGCGCTGCGAGGCGCTGCGGCGATGCGGCTTCTGAGAGAC 790

QY 481 GCGGTGAATATGCAACAGGAAATTTACCGGTTGCTCTTCTCATCTCTCTGAGCT 540  
DB 791 GCGGTGAATATGCAACAGGAAATTTGCGCGTTGCTCTTCTCATCTCTCTCTAGCT 850  
QY 541 TTGCTGCTCTGTTTACCATTCAGCTTCCGCTTATGAAATGCGCAAGTGTCCGGATC 600  
DB 851 TTGCTGCTCTGTTTACCATTCAGCTTCCGCTTATGAAATGCGCAAGTGTCCGGATC 910  
QY 601 TACCATGTCAGAAAGATTTGGCTCCCACTCAAGCATCGGTGCGAGACGGGACATGTC 660  
DB 911 TACCATGTCAGAAAGATTTGGCTCCCACTCAAGCATCGGTGCGAGACGGGACATGTC 970  
QY 661 ATGCAACACCCCGGCTGTGTGCTCTGTCGCGAGGTAATTCCTCCGCTGCTGGGTA 720  
DB 971 ATGCAACACCCCGGCTGTGTGCTCTGTCGCGAGGTAATTCCTCCGCTGCTGGGTA 1030  
QY 721 GCGCTCACTCCACGCTGCGCGCCAGAGACCGCAGATCCCACTGCGCAATACAGCG 780  
DB 1031 GCGCTCACTCCACGCTGCGCGCCAGAGACCGCAGATCCCACTGCGCAATACAGCG 1090  
QY 781 CACGTGATTTGCTGCTGCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 1091 CACGTGATTTGCTGCTGCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150  
QY 841 TCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
DB 1151 TCGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210  
QY 901 GTACAGACCTGCAATTTGTTCAATCTATCCCGGCAAGTATGATGATGATGATGATGAT 960  
DB 1211 GTACAGACCTGCAATTTGTTCAATCTATCCCGGCAAGTATGATGATGATGATGATGAT 1270  
QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
DB 1271 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1330  
QY 1021 CCACAGCTGTCGTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
DB 1331 CCACAGCTGTCGTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1390  
QY 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
DB 1391 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1450  
QY 1141 GTTGAACGGGAAACCTTACAGCAAGGGGAGACACAGCGCGCGCCAGCGGCTTACA 1200  
DB 1451 GTTGAACGGGAAACCTTACAGCAAGGGGAGAGATGATGATGATGATGATGATGATGAT 1510  
QY 1201 TCCCTCTTCAACCTGAGCGCGCTCAGAAATTCAGCTTGTAAACACCAACGCGAGCTG 1260  
DB 1511 ACCCTCTTATCATGAGGCTGCTCAGAAATTCAGCTTGTAAATACCAATGGGAGCTG 1570  
QY 1261 CACATCAACAGAACTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
DB 1571 CACATCAACAGAACTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1630  
QY 1321 CTGTTCTTACAGCAAGGCTTCAATGCGTCCGGAATGCTCAGAGCGGATGCGAGCTGCG 1380  
DB 1631 CTGTTCTTACAGCAAGGCTTCAATGCGTCCGGAATGCTCAGAGCGGATGCGAGCTGCG 1690  
QY 1381 CCATTTGACAGATTGATCAGAGGCTGAGGCTCCATCACTTAAATGAGTCCACGCGCTTG 1440  
DB 1691 CCATTTGACAGATTGATCAGAGGCTGAGGCTCCATCACTTAAATGAGTCCACGCGCTTG 1750  
QY 1441 GACCAAGGCTTATTTGCTGCGACTAGCGCACTTAACCTGTGTGATGATGATGATGATGAT 1500  
DB 1751 GACCAAGGCTTATTTGCTGCGACTAGCGCACTTAACCTGTGTGATGATGATGATGATGAT 1810  
QY 1501 CAGGTGTGCGCCAGGCTATCTGTTTCACTCAAGCCCTGTTGTGTGGGGACAGCGAT 1560  
DB 1811 CAGGTGTGCGCCAGGCTATCTGTTTCACTCAAGCCCTGTTGTGTGGGGACAGCGAT 1870  
QY 1561 CGTTTGGCGCCCTTACGATAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1620

```

Db      1871 CGTTCCGGGGCCCCCGACGTAACCTTGGGGGAGATGAGACGACCTGCTACTTCTTCAAC 1930
Qy      1621 AACACGCGGCGCCGACGAGGCACTGTTGCGCTGTAACATGATGAAATAGACCGGGGTTG 1680
Db      1931 AACACGCGGCGCCGACGAGGCACTGTTGCGCTGTAACATGATGAAATAGACCGGGGTTG 1990
Qy      1681 ACCAAGAGCTGTGGGGGGCCCCCGTGCACATCTGGGGGGGTGGGCAACAACCTTTGATC 1740
Db      1991 ACCAAGAGCTGTGGGGGGCCCCCGTGCACATCTGGGGGGGTGGGCAACAACCTTTGATC 2050
Qy      1741 TGGCCCGACGAGCTGTTCCGGAAGACATCCCGAGGCCACTTACCAAAATCGGTTGGGG 1800
Db      2051 TGGCCCGACGAGCTGTTCCGGAAGACATCCCGAGGCCACTTACCAAAATCGGTTGGGG 2110
Qy      1801 CCTGGTTG 1809
Db      2111 CCTGGTTG 2119

RESULT 9
AF207761 9379 bp RNA linear VRL 27-Apr-2000
LOCUS Hepatitis C virus strain MD20 complete genome.
DEFINITION AF207761
ACCESSION AF207761 GI:7650239
VERSION AF207761.1 GI:7650239
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (bases 1 to 9379)
AUTHORS Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Marumo,F. and Sato,C.
TITLE Characteristics of hepatitis C viral genome associated with disease progression
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 9379)
AUTHORS Nagayama,K., Kurosaki,M., Enomoto,N., Miyasaka,Y., Izumi,N. and Sato,C.
TITLE Direct Submission
JOURNAL Submitted (23-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan
FEATURES
source 1. 9379
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="MD20"
/db_xref="taxon:11103"
/notes="Isolated from a patient with persistently-normal ALT values (ASG-12)"
ALT values (ASG-12)
330. .9362
/codon_start=1
/product="polyprotein"
/protein_id="AF65951.1"
/db_xref="GI:7650240"
/translation="MSTNPKPQRTKNTNRPDVYKFGGGOIVGGVLLPRGRPL
GYATRTGSRQPRROPVPRARQEGMAAPGPMPLVNEGMAAGMLSPRG
SRPMGPTDPRRSRLKIVITLTCGSPADIMGYIPVAGPLGAAALAHGRVLD
GVNVAQNLPGCSFSLFALLSLCLTPASAYVRVSVGYOVTNCSNATYVEAD
MIMHTGCTPCVARENSSRCWALLTFLAARASITPTTYRRVLDLVGAACSAHY
VDLCSVFLVSQLFTFSRRHETVDDNCSTIPGVHTGRHAWMMNWSPTAALV
SOLRLPQAVVNDVAGAHWGLAGLAYSVWVMAVLLVMLFAGVDGNTAALV
ASTTGAVLSFRSGANQIKQIINTNGSMHNTALNCNBSINTGMLAALYTRKFNAS
GCEPLAHGCSIDQPAQMGPIHTHTVPASDORPCMHVAPRCGIVPASOVCGPYVC
FPSPVVGTTDTRFGVPTVWGENETDVLINNSRPLNPGCTMMNNSGPKTGGG
PCPNIGVGNTLTCTPDRCRKPRTATYTCGSGPMLTRPCMHYKRLMNYCTNPF
TTRKRVYGGVEHRLNACNMTGRCEDLDRKSLPLSLSTTBMOVLPCSTTL
PAIATGILHLHONIIVQOYLUGISVVVSAIKWEYIILFLILADARVACIMMML
IAQBAALBNLVLINAAVAGAGTSLFVPCAAVYIKRLIPGAAYAVGVPWPLL
LILALPRAVAMREMAAGSCGAVFTGLALTLSPHYKVLARLIMWLYOFSRAEL
LQWIRPLANVGRDAIILLTCLVHBEILFTIKILLAVGLPMLQAGITKIVFYVR

```

## ORIGIN

```

Query Match      87.4%; Score 1580.2; DB 14; Length 9379;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1666; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy      1 ATGAGCAGAAATCTTAACCTCAAGAAAGAAACCAACGTAACCAACCGCGCCGACAG 60
Db      330 ATGAGCAGAAATCTTAACCTCAAGAAAGAAACCAACGTAACCAACCGCGCCGACAG 389
Qy      61 GACGCAAGTCCCGGCGGAGTGGTCAAGTCTGTTGAGATTAACTGTTGCCGCGCAG 120
Db      390 GACGCAAGTCCCGGCGGAGTGGTCAAGTCTGTTGAGATTAACTGTTGCCGCGCAG 449
Qy      121 GGGCCGAGTGGGTGGCGCGGACATGAGAAAGACTTCCGAGCGGTGCAACTCGTGA 180
Db      450 GGGCCGAGTGGGTGGCGCGGACATGAGAAAGACTTCCGAGCGGTGCAACTCGCGA 509
Qy      181 AGCGCAACCTATCCCAAGGTCGCCAGCCCGGAGGAGCGGCTGACGCCCGG 240
Db      510 AGCGCAACCTATCCCAAGGTCGCCAGCCCGGAGGAGCGGCTGACGCCCGG 569
Qy      241 TACGCTGGCCCTCTATGAGCAAGAGGAGTGGGTCAGAGATGAGTCTCTGACCC 300
Db      570 TACGCTGGCCCTCTATGAGCAAGAGGAGTGGGTCAGAGATGAGTCTCTGACCC 629
Qy      301 CGTGAGTCCCGGAGTGGGAGCCCACTGACGACCCCGGAGTGGTGGGTAATTTGGG 360
Db      630 CGTGAGTCCCGGAGTGGGAGCCCACTGACGACCCCGGAGTGGTGGGTAATTTGGG 689
Qy      361 AAGCATGATGATCCCTGACATGCGGCTTGGCGGACCTGATGGGGTAACTTCGCTGTC 420
Db      690 AAGCATGATGATCCCTGACATGCGGCTTGGCGGACCTGATGGGGTAACTTCGCTGTC 749
Qy      421 GGGGCTCCCTTGGGGGCGCTGCGAGGGCCCTGCGCAGTCCGGGTTCTGAGAGAC 480

```

```

AHGILRACMLVRKVGAGHYLQMAFMKLAALTSTVYVYDHLPLRDMWHAHGLRDLAAVE
PVPFSGDMERKILTWKADTAAGDIIISGLPVSARBRREILLAGPDSLGGGMYHLGPK
AYAOOTRGLGCIITSLTRGRDNQVQBEVQVYSTAQSLATACNVCNCTVYHAGSGK
ALAQOKPVTQMTVNDVQVQAPPGKASLTPTCGSSDLYLTRADVIVPRRG
DSRGSILSPRPVSLKSGSGLPCPSGAVIIPRAVCTGRVAAVADVPEVSEMTT
MRSPIFTDMSPPAVQFQVVAHLAPTSQSGTKVPAVAAQGVVLAIVLNSVATL
GEGVANSKAGVDPNRTGRTITTTGAATPTSTYGVFLADGGCGGAGAYIILICDEBS
TDSITLIGITVLDLDAETGARTVLVLAATTPPGSVVPHNIEVLAISNTGELPEFK
AIPETIKGRLILCHSKKKDELAKSLAAGLNAVAAYRGLDVSVITPSGDVVVVA
TDLMTLFTGDDVSVIDCNTVQIVDSLDFTLTLETITVEQDVAVSQORRRTGRG
RMGIYRFPVPRPSGMPDSVLCBCYDAGCMYELTPAETSVRLAYLNTPGLPVCO
DLHEPESVFTGLTIDAFISQTOAGNPFYLAQVAVCAAPAPPSDMMK
LRLKPTLHGPFLYLRGAVONEVLTHTPIKYIMACMSADLEVSTSMVVGVLTA
ALAAVCTTGTCVIVIGRIISGRPAVIPRODLVQEPDEMERASHLPIEBOGMLAB
QKOKRALGLQATATQAEBAAPVSKRALETTPAKKMERATISGIVYAGLTPGN
PAIASLMAFTASITSPLTQHTLLFNILGAGVAQVAPPSASAPVAGACIAGAAGSI
GLGKVLVDLAGYAGVAGVLAIVKMSGEMSTBDLVNLAIPALISPGLVGVVCAAI
ILRRVYGPBGAVQMMNRLIAPASGKNVSPHYVPESDAARVQIILSLITDOLK
RLHOMINEDCSTPGSGMLRDVMDVCAVLTPDYKTMOSKLLPLRGVFPFCGSGYK
GIWRGDGIMHTTCCPGAQITGVHNGSMRIVGPKCSNTMHTGTPPINAVTTPCPSP
APNYSRALMRVAAEVEYTRVGDHYVYGTMTDNYVKQCOVPAEPFEDGILHR
YAPACKPLREBEVTFQVGLNQLYVGSQULPCEBEPDVAVITSLMTPDSHTAETARRL
ARGSPSLASSASQLSASLAKATCTTRHSDPADLITENILMRQEMGNITRVESEN
KVIILDSFELQAEDEBRESIYAAELIKSRKFPAMPVPAAPADVPNPLBESKQEDY
VPPVHGCPLPTKAPPIPPPRKRTVLTSTVSALALATKTFGSESSAVIDSGT
ATAPDQASGDGAGSDVESVSMPLBEGEPDPLSDGMSGVSEASBEDVVCMS
YTMGTALITPPCAESKLPINLMSLSLRHNMVYATTSRSASQOKKYTPRMOVLD
DHYRDVLEMKRAKASTVAKSLSVBACULTPPHSRKFSFGAKOVNLSKATHT
RSVMQDLADESEPTIDTYMANNEVCFQPEKGRPALIVPDLGVSEKEMALYD
VSTLPLQAVMSGYSGFQYSPQGRVEFLVVAASKXPMFAADVTRFDSVENDRIV
ESITVOCCDLAPEAROALISLTERLYIGPFLNSGONCGYRCSAGSLTSCGTI
TCYIKASACRAKAKODCTMLYCGDDIIVYGSAGTGOEASLRYVTEMTYTSAPRG
DPPQPEYDLIELTSCSNYSVVAHDSGRVYVLTDPPTPLARABETARHPNYSKL
GASIMVAPLTMARMLLMTHTFSIILAQOELRALDQITGACTSLBEPDPIOTLKH
GLASISLHSGPGEINRVASCRLKIVPLRWRHRAVRKALLISQGRAPATCGXYL
FMVAATKTKLTPIPAASHLDLSGWFVAGSGDIYHSLRRARPRMLMCLLLSVGV
GYLLPNR"

```

Db 750 GGGCCCCCTTAGGGGGCGCTGCGACAGGCGCTGCGCATGCGGCTCCGGGTTCTGAGAGAC 809  
Qy 481 GCGTGAATATAGCAACAGAGAAATTACCGGGTGGCTTTCTATCTTCCCTTGGCT 540  
Db 810 GGGGTAACTATAGCAACAGAGAAATTGCCCCGTTTCTTCTATCTTCCCTTGGCT 869  
Qy 541 TTGCTGCTGTTTGAACCAATCCAGCTTCGCTTATGAGTGGCAACGTCGCGGATC 600  
Db 870 TTGCTGCTGCTTGAACCAATCCAGCTTCGCGTATGAGTGGCAACGTCGCGGATC 929  
Qy 601 TACATATAGCAACGATTTGCTCAACTCAAGATGTTATGAGACAGCGCATGATC 660  
Db 930 TATCAGGTCACGAACGATGCTTCAACGACATTTGTATGAGGCAACGATGATC 989  
Qy 661 ATGCAACACCCCGGGGTGTGCGCTGTCTCCGGAGAGGTAATTCCTCCGCTGTGGGTA 720  
Db 990 ATGCAACACCCCGGGGTGTGCGCTGTCTCCGGAGAGATTAATCTCTCCGCTGTGGGTA 1049  
Qy 721 GGGCTACTCCCAAGCTGCGGGCAAGAGCGCAGCATCCCACTGCGCAATAGACGC 780  
Db 1050 GGGCTACTCCCAAGCTGCGGGCAAGAGCGCTAGATCCCAACGACGATGGACGT 1109  
Qy 781 CAGCTGATTTGCTGTTGGGGCGGCTGCTTCTGCTCCGCTATGATGCTGGGGATCTC 840  
Db 1110 CAGCTGATTTGCTGTTGGGGCGGCTGCTTCTGCTCCGCTATGATGCTGGGGATCTT 1169  
Qy 841 TCGGATCTGTTTCTCGCTCTCAGCTGTTCACCTTCGCTGCGCGCATCATGACAG 900  
Db 1170 TCGGATCTGTTTCTCGCTCTCAGCTGTTCACCTTCGCTGCGCGCATCATGACAG 1229  
Qy 901 GTACAGAGATGCAATTTGTTCAATCTATCCCGGCAAGTACAGGTCAACCGCATGCTTG 960  
Db 1230 GTACAGAGATGCAATTTGTTCAATCTATCCCGGCAAGTACAGGTCAACCGCATGCTTG 1289  
Qy 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
Db 1290 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349  
Qy 1021 CCACAGAGCTGTGTTGGAATGTTGGAATGTTGGAATGTTGGAATGTTGGAATGTTG 1080  
Db 1350 CCACAGAGCTGTGTTGGAATGTTGGAATGTTGGAATGTTGGAATGTTGGAATGTTG 1409  
Qy 1081 TACTATTCATGTTGGGGAATGTTGGGGAATGTTGGGGAATGTTGGGGAATGTTGG 1140  
Db 1410 TACTATTCATGTTGGGGAATGTTGGGGAATGTTGGGGAATGTTGGGGAATGTTGG 1469  
Qy 1141 GTTGAACGGGGAATCTTACACGACAGGGGGGAACACAGCGCGCGCCCAACGGGCTTACA 1200  
Db 1470 GTTGAACGGGGAATCTTACACGACAGGGGGGAATCTTACACGCGCGCGCTTGTG 1529  
Qy 1201 TCCCTTTTACACCTGCGCGGCTCAGAAATTCAGCTTGTAAACCAACGACGAGCTTG 1260  
Db 1530 TCTCTTTTACACCTGCGCGGCTCAGAAATTCAGCTTGTAAACCAACGACGAGCTTG 1589  
Qy 1261 CACATCAACAGAACTGCTTGAATGCAATGATCTCCCTTCAATGAGGTTCTTGGCGCG 1320  
Db 1590 CACATCAACAGAACTGCTTGAATGCAATGATCTCCCTTCAATGAGGTTCTTGGCGCG 1649  
Qy 1321 CTGTTCTTACACCGCAAGTTCAATGCTCCGATGCTCAGACCGCATGCGCAGCTGCGC 1380  
Db 1650 CTGTTCTTACACCGCAAGTTCAATGCTCCGATGCTCAGACCGCATGCGCAGCTGCGC 1709  
Qy 1381 CCCATTTGACAGATTCAGAGGGGTGGGTTCCCATCACTTATATGATGCCACGAGCTTG 1440  
Db 1710 TCCATTTGATCAATTCGCCCGAGGGGTGGGTTCCCATCACTTATCACTCACTGTGCCGCGCAGCTTG 1769  
Qy 1441 GACGAGAGGCTTATTTGCTGGAACTAACGACCTCAACCGTGTGATTCGTGCGCGCTTG 1500  
Db 1770 GACGAGAGGCTTATTTGCTGGAACTAACGACCTCAACCGTGTGATTCGTGCGCGCTTG 1829  
Qy 1501 CAGGCTGTGCGCGCAGTGTACTGTTTCACTCCAGCGCTGTTGTTGTTGGGAGCAACCGAT 1560  
Db 1830 CAGGCTGTGCGCGCAGTGTACTGTTTCACTCCAGCGCTGTTGTTGTTGGGAGCAACCGAT 1889

Qy 1561 GCTTTCGGCGCCCTTACAGATGAGGTTGAGATGAGACGAGTGTCTTCAAC 1620  
Db 1890 GCTTTCGGCGCTTCTTACAGATGAGGTTGAGATGAGACGAGTGTCTTCAAC 1949  
Qy 1621 AACACGCGCGCGCAACGAGGGAATGTTGAGTGTGATGATGATGATGATGATGATG 1680  
Db 1950 AACTGCGCGCGCGCTGAGCACTGTTGAGTGTGATGATGATGATGATGATGATG 2009  
Qy 1681 ACCAAGAGTGTGAGGCGCGCGCGCGCAACCTGAGGAGGTTGAGCAACCTTGTATC 1740  
Db 2010 ACCAAGAGTGTGAGGCGCGCGCGCGCAACCTGAGGAGGTTGAGCAACCTTGTATC 2069  
Qy 1741 TGCCCCAGAGCTGTTCCGAGAGATCCAGAGCACTTACACCAATGCGGTTGCGGG 1800  
Db 2070 TGCCCCAGAGTGTGTTCCGAGAGCACTTACAGCACTTACACCAATGCGGTTG 2129  
Qy 1801 CTTTGGTTG 1809  
Db 2130 CTTTGGCTG 2138

RESULT 10  
AF207766  
LOCUS AF207766 9379 bp RNA linear VRL 27-APR-2000  
DEFINITION Hepatitis C virus strain MD25 complete genome.  
ACCESSION AF207766  
VERSION AF207766.1 GI:7650249  
KEYWORDS  
SOURCE Hepatitis C virus  
ORGANISM Hepatitis C virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
REFERENCE 1 (bases 1 to 9379)  
Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and Sato, C.  
Characteristics of hepatitis C viral genome associated with disease progression  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 9379)  
Nagayama, K., Kurosaki, M., Enomoto, N., Miyasaka, Y., Izumi, N. and Sato, C.  
Direct Submission  
JOURNAL Submitted (23-NOV-1999) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-ku, Tokyo 113-8519, Japan

FEATURES  
source  
CDS  
1..9379  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/strain="MD25"  
/db\_xref="taxon:11103"  
/note="Isolated from a patient with liver cirrhosis and hepatocellular carcinoma (HCC-5)"  
330..9362  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="AA05956.1"  
/db\_xref="GI:7650250"  
/translation="MSTNPKRQRTKENTNRPODVKPPGGQIVGAVLLPRGRPLGVARTRTISERSQGRGRQPIPKARPEGRTWAPQVWPVYEGMGAMGMLSPRGSRPMSGPTDPPRRSRNLGKVIDLTGFSADLMGYIPIVGAFLGGAALAHAGVLEBGNVATGNLPCSSISFLIALISCLTIPASAYEVNAGVHVNDCNSIYVEADMITKTPGCVPRERNSSRCVVALPTLTAARNAVPTKTRRHVDLLVGAAPGSAVYVGDICGVFLVSQLPFTSPRRHETVQDNCSTYEGHLSGHRMADMMNNSPTALIVSOLLRIQAVMDVWAGAHWGLAAYSMVGNAAKXVLIWLLPAGVDSYRVGGTAGRITVGLARFPTSGPSONIQVNTNGSMHINRTLANCDSIQTFIALFTYHRENDSGPERMASRSRIDKFEQMGPIYAGDSLDORPYCWHYAPQPGIYPAAGVCPVYCFPSPIVVVGTSDRSRGVPTVMGNVESVLLLNMRPQGMNFGCTPMYNSGTFTCCGPPCNIGVGNNTLLCPDPCRKHBEARYTICGSSPWLTPRCLVDYPTRLMYPTCVNFTTLKRVNYGGERLNAACMWTREGCNEDRDRSELPLISTEMQILLPCSFITLIPALSTGLIHLQNVNDVQVLYGISAVSVFKIKWEYVLLFLIADARVACILMMILLIAQEAALIEKLIVINAASVAGAHGILLFLVFCFAAWYIKKLIVGAAVYALYGVPLLL

LLALPELAAMREMAASCGAVFGLVLLTUSPHYKFLARLIMWLQVLTREAL  
 LOYVPEPLNRGRDAILILLCAHBEPLIDISKILLATIGLPICTVAGTTRIPVPR  
 AGLIRVCHLVKVAAGHYQYOMALMKLALITVYDHLTPLEDMAGHRLAIVAVE  
 PVLPSDEKTVITWADTAQVNIISGLPVSAKREKILLEGPARDEHGRWRLAAPT  
 AYQOQRLGLGCIITSLTGRDNQVGEVAVSTATOSPLATCVNGCVVYAGASK  
 TLGAKPITOMYTNVODLVMLVPAEVAEVLPTCGSSDLTVLRHADVILPVRG  
 DSGSLSPRPVYILKSGSGGLCPGSHGVGIFRAVCTRGAKAVDPVPMEMT  
 TRSPPTDSSPAVPOTPQVAILHAPTSGKSTKVPAAVAGVAVIVPVAATL  
 GRAYMSKAGVDPNIRITGVRTTTGAPITSTYTGFLADGCGSGAYDINKDECHS  
 TBSITSLIGITVLDAVAGARLVATAPGSVTVPHNIEILSNSTGELIPFYK  
 ALPIEIKGRHILFCHSKKCKDELAASKGLGNVAVYRGADVITPVGSVVVA  
 TALMTGPTGDLSDVDCNTCVQTFDFSLPFIETTTETTPDASRSRORRTEG  
 RRIYRFTVPERSGMPSVCECDACAVELTPATSVRLAAYLNTPELPHYCO  
 DHEFESVFTGLTHIDAHLSOTKQAGNMLPYLVAOATASADILEVTVLVGCTLA  
 LIRKPTLSPPTLIRLGAONEVILITHTYIMACNSADILEVTVLVGCTLA  
 ALAAVCLTGSVVIVGRILISGRPAVDEVLVREFDEBECAASHLPITEQMOAB  
 OFKOKALGLQTSKQAEAAVAVESKRWLEAFMAKHMMNFTSGOYLAGLSTLGN  
 PAIASLMAFTASITSPLTQSTLLPILGGMVAALPAPASAFVAGIAGAVASI  
 GIKVAVDLIAGVAGVAGLVAFKVMGSEMPSTEDLVNMLPALISPGALVGVGSI  
 ILRBHYGBEGAVQMMNRILAPASRGNHSPTHYVPSDAAVAVTOLISLITTLQK  
 RLMQINEDSTPCSGSWLMDVMDITCTVLTRKTLWQSLFLRLGVPPLSCQKRGK  
 GVMRGDGIHMTTCPCGAQITGHVKNQSMRIVGRTCSNTHGTFPVNAYITGCTSP  
 APVYSKALMRVAEEVAVTVGDVPHYVTGMDTVKCPQVPAPEFTEVDVRLHR  
 YAPACKPLREEVTFQVGLNOYLVGSQLPCEPEPDVAVLTMLDPSHTAETAKRL  
 AAGSPSLASSASOLASBLKATCTTRHDSPPADLILENMLRORMGNITRVESN  
 KYIILDSPEPLRAEBDEBREVSAEVLRSRKEPPALPIYARDPNPLLESKEDY  
 IIPVHVGCPLEPTKAPPIPEPRKRKTIVLSTVSSALBWLAKRTGSESSSVAVDGS  
 ATAPPOPDGADGSDVESISMPLEBGPDPDLSDASMTVSEBASDSSVCMS  
 YMTGALITPCAAESKLPINALNSLILHNNHVAATTSASQORQKTFDLQVLD  
 DHYRGVLEIKAKASTYKALKSVESAKLTPHARSKSYKAGVAKVNLISGAVNHI  
 RSVMKDLLEDETETPIDITTIMAKNEVFCVQPEKGRKPARLIVPDLGIVCEMAYD  
 VYSTLPOAWGSSYGFQYSPQGRVEFLVNAKSKRPMGCSYTRCPDVSITENDRT  
 EBSITVCCDLAPBARQAIRSLERLYVGGELTNSKQNGCYTRCARSPVYTESGNTL  
 TCYKASACRAKILQDCTMLVCGDVLVYCEBAGTQEBASLSTVTEAMRTASBP  
 DPQPEYDLELITSCSNVSVADHAGSKRYVLTTRDPTPLAAAEVTRASTVSNML  
 GNIMYAPTLMAIMILMTFSLILAOBELKALDCOIGAICSIPELDPOLIOILH  
 GLSASFILSYSPGEINRVAISLCKRLGVPRLVRRLRARSVACSLISQGRRAATCGYL  
 FNVAVTKLKTPIPAASQDLDSNMFVAGISGDIYHLSRARPFRMCLLILSIVG  
 GYLLENR

## ORIGIN

Query Match 87.4%; Score 1580.2; DB 14; Length 9379;  
 Best Local Similarity 92.1%; Pred. No. 0;  
 Matches 1666; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 1 ATAGACGAATCTTAAACCTCAAGAAAAAACAACGTAACAACAACCGCGCCACAG 60  
 DB 330 ATGAGCAGCAATCTTAAACCTCAAGAAAAAACAACGTAACAACAACCGCGCCACAG 389

QY 61 GAGCTCAAGTCCCGGGCGGTGTCAAGTGTGTGTGAAGTTTAACTGTTCGCGCGAG 120  
 DB 390 GAGGTAAAGTCCCGGGCGGTGTCAAGTGTGTGTGAAGTTTAACTGTTCGCGCGAG 449

QY 121 GGGCCAGGTTGGGTGTGCGCGAGTAGGAAGACTTCGAGCGGTGCAACCTCGTGA 180  
 DB 450 GGGCCAGGTTGGGTGTGCGCGAGTAGGAAGACTTCGAGCGGTGCAACCTCGTGA 509

QY 181 AGGCGCAACCTATCCCAAGGCTGCGACGCGAGGCGAGGCGTGGGCTGAGCCCGG 240  
 DB 510 AGGCGCAACCTATCCCAAGGCTGCGACGCGAGGCGAGGCGTGGGCTGAGCCCGG 569

QY 241 TACCTTGGCCCTCTATGGCAACGAGGGATAGGGGTGGGAGAGATGGCTCTGTCA 300  
 DB 570 TACCTTGGCCCTCTATGGCAACGAGGGATAGGGGTGGGAGAGATGGCTCTGTCA 629

QY 301 CGTGGCTCCGGGCTAGTGGGCGCCCACTGACCCCGGCGTAGTTCGCTAATTTGGGT 360  
 DB 630 CGTGGCTCCGGGCTAGTGGGCGCCCACTGACCCCGGCGTAGTTCGCTAATTTGGGT 689

QY 361 AAAAGCATGATACCTTCACATGCGGCTTTCGCGACCTCATGGGGATCATTCGCTGTC 420  
 DB 690 AAGGTATGATACCTTCACATGCGGCTTTCGCGACCTCATGGGGATCATTCGCTGTC 749

QY 421 GGGCTCCCTTAGGGGGCGCTGCGAGGGCCCTGGCGCATGGGCTCCGGGTTCTGAGGAC 480  
 DB 750 GGGCCCCCTTAGAGAGCGCTGCCAGAGCCCTGGCGCATGGGCTCCGGGTTCTGAGGAC 809

QY 481 GGGGTAACTATGCAACAGGGAAATTTAACCGGTTGCTTTCTATCTTCTCTTGGCT 540  
 DB 810 GGGGTAACTATGCAACAGGGAAATTTAACCGGTTGCTTTCTATCTTCTCTTGGCT 869

QY 541 TTGCTGTCTGTTTACCAATTCAGCTTCGCTTATGAAAGTGGCAAGCTGTCCGGATC 600  
 DB 870 CTGCTGTCTGTTTACCAATTCAGCTTCGCTTATGAAAGTGGCAAGCTGTCCGGATC 929

QY 601 TACCATGTACGAACGATTTGCTCAACTCAAGCATGTGTGACGACAGCGACATGATC 660  
 DB 930 TACCATGTACGAACGATTTGCTCAACTCAAGCATGTGTGACGACAGCGACATGATC 989

QY 661 ATGCAACCCCGGGGTGTGCTGCTGTCGCGAGGGTAAATTCCTCCGCTGCTGGGTA 720  
 DB 990 ATGCAACCCCGGGGTGTGCTGCTGTCGCGAGGAAACATCTCCGTTGCTGGGTG 1049

QY 721 GGGCTACTCCACGCTCGCGGCAAGACGCCAGATCCCATGCGACATACGACGC 780  
 DB 1050 GGGCTACTCCACGCTCGCGGCAAGACGCCAGATCCCATGCGACATACGACGC 1109

QY 781 CAGCTGATTTGCTGTTGGGGGCGGCTCTGCTCCGCTATGATAGTGGGGATCTTC 840  
 DB 1110 CAGCTGATTTGCTGTTGGGGGCGGCTCTGCTCTGCTATGATAGTGGGGATCTTC 1169

QY 841 TGGGATCTGTTTCTCTCTCTCTAGCTTTCAACCTTCGCGCTGCGGACATACGAGC 900  
 DB 1170 TGGGATCTGTTTCTCTCTCTCTCTAGCTTTCAACCTTCGCGCTGCGGACATACGAGC 1229

QY 901 GTCAGAGATGCAATTTGCTCAATCTATCCCGGCAAGTACAGGTACCGCATGCTTGG 960  
 DB 1230 GTCAGAGATGCAATTTGCTCAATCTATCCCGGCAAGTACAGGTACCGCATGCTTGG 1289

QY 961 GATATATATGATGATCTGTCACTTACAGAGCCCTAGTGTATCGAGTACTCGGATC 1020  
 DB 1290 GATATATATGATGATCTGTCACTTACAGAGCCCTAGTGTATCGAGTACTCGGATC 1349

QY 1021 CCACAAGCTGTGCTGACATGTGTGCGGGGGCCCACTGGGAGATCTTGGCGGCTTGGC 1080  
 DB 1350 CCACAAGCTGTGCTGACATGTGTGCGGGGGCCCACTGGGAGATCTTGGCGGCTTGGC 1409

QY 1081 TACTATTCATGAGTGGGAACTGGGCTAAGTCTTGTGTGATGCTACTTTCGCGGC 1140  
 DB 1410 TACTATTCATGAGTGGGAACTGGGCTAAGTCTTGTGTGATGCTACTTTCGCGGC 1469

QY 1141 GTTGAAGGGGAACCTTACACGACAGGGGGGACACACGGCGCGGCCCAACGGGCTTAA 1200  
 DB 1470 GTTGAAGGGGAACCTTACACGACAGGGGGGACACACGGCGCGGCCCAACGGGCTTAA 1529

QY 1201 TCCCTTTTCAACCTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACAGGCACTGG 1260  
 DB 1530 CCGTTCTTACATCTGGGCGGCTCAGAAATTCAGCTTGTAAACCAACAGGCACTGG 1589

QY 1261 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCAAACTGGGTTCTTTCGCGC 1320  
 DB 1590 CACATCAACAGAACTGCTTGAATGCAATGACTCCCTCAAACTGGGTTCTTTCGCGC 1649

QY 1321 CTGTTCTTACACCAAGTTTCAATGCTGCGGATGCTAGACGATGCGACAGCTGGCCG 1380  
 DB 1650 CTGTTCTTACACCAAGTTTCAATGCTGCGGATGCTAGACGATGCGACAGCTGGCCG 1709

QY 1381 CCCATTGACCAATTCATGAGGGGTGGGCTCCCATCATCTTATATATGATCCACGGCTTG 1440  
 DB 1710 TCTATTGACCAATTCATGAGGGGTGGGCTCCCATCATCTTATATATGATCCACGGCTTG 1769

QY 1441 GACCAAGGCGCTTATGCTGCGCATACGCACTCAACGCTGTGTATGCTGCGCGCTTG 1500  
 DB 1770 GACCAAGGCGCTTATGCTGCGCATACGCACTCAACGCTGTGTATGCTGCGCGCG 1829

QY 1501 CAGGTGTGTGGCCAGTGTACTGTTTCACTCAAGCCCTGTTGTGGTGGGAGACACCGAT 1560



|||||  
1830 CAGGTGTGTCAGTGTATGCTTCAACCCCAAGCCCTGTTGTGTGGGACGAGTAT 1889  
|||  
1561 CTTTCGGGGCCCCCTACGACATGAGGTGAGATGAGACGAGCTGCTTCTCAAC 1620  
|||  
1890 CCGTCGGGGCTCCCTACGATATACGAGGGGGGAAATGAGTCCGACGTCTCTCAAC 1949  
|||  
1621 AACACGGCGCGCCGACGAGGGGCAACTGCTGCGGCTGTACATGATGATGATGACCGGGTTC 1680  
|||  
1950 AACACGGCGCGCCGACGAGGGGCAACTGCTGCGGCTGTACATGATGATGATGACCGGGTTC 2009  
|||  
1681 ACCAAGACGTGTGGGGGCCCCCGTGTCAACATCGGGGGGTGCGCAACAACCTTTGATC 1740  
|||  
2010 ACCAAGACGTGTGGGGGCCCCCGTGTCAACATCGGGGGGTGCGCAACAACCTTTGATC 2069  
|||  
1741 TGCCCCAGGACGTGTCTTCGCGGAAGACATCCCGAGGCGACCTTAACCAAAAACGGGTTTCCGGG 1800  
|||  
2070 TGCCCCAGGACGTGTCTTCGCGGAAGACATCCCGAGGCGACCTTAACCAAAAACGGGTTTCCGGG 2129  
|||  
1801 CCTGTGTTG 1809  
|||  
2130 CCTGTGTTG 2138  
|||  
RESULT 11 9547 bp RNA linear VRL 22-AUG-2002  
AB049091  
LOCUS Hepatitis C virus gene for polyprotein, complete cds,  
DEFINITION isolate:HCVT142.  
ACCESSION AB049091.1 GI:11559448  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Hepatitis C virus  
Hepatitis C virus  
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepadnavirus.  
REFERENCE  
AUTHORS  
1  
Takahashi, K., Iwata, K., Matsumoto, M., Matsumoto, H., Nakao, K.,  
Hachihara, T., Ohta, Y., Kanai, K., Maruo, H., Baba, K., Hijikata, M. and  
Mishiro, S.  
Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients  
with hepatocellular carcinoma: the 'progression score' revisited  
Hepatol. Res. 20 (2), 161-171 (2001)  
11348851  
2 (bases 1 to 9547)  
Mishiro, S.  
Direct Submision  
Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,  
Department of Medical Sciences, 6-3-22 Higashi Oh-1, Shinagawa-ku,  
Tokyo 140-8522, Japan (E-mail: shunji.mishiro@po.toshiba.co.jp,  
Tel:81-3-3764-8981, Fax:81-3-3764-8992)  
FEATURES  
source  
1..9547  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/isolate="HCVT142"  
/specific\_host="human"  
/db\_xref="taxon:11103"  
/tissue\_type="serum"  
/note="HCV-genotype: 1b"  
<1..259  
/note="partial"  
260..9292  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="BAB18804.1"  
/db\_xref="GI:11559449"  
/translation="MSTNPKPQKTKRNTNRRPDVKEFGGQIVGVYLLPPRPRL  
GVRATKTSERSGPRGRQPIKARQPEGRAMAQGPMPPLVNGSLGMAAGMLSPRG  
SRBSGPTDPRRSRNIGKVIDITLCGFADLMGYIPIVAGAPICGAARAKLAHGVRLPD  
GVNATGNLPGCSFSIFLLALISCLTIPASATVENVSGITHVTNDCSSSI VYEAD  
VTMAHAGCVPCVRNNSSRCWALPTTLAARNSVPTTIRRHVDLLVGTAAFCSAMV  
VDLDCGSVELISQLFTFSRHRHETVQSCNSLYPGHIVGHRMAWMMNWSPTALVY

SOLLRIDPQAVMDVAGAMGVLGAGIAYYVNMNAKYLIVMLFAGVDDGTHVTGAGQ  
GSAFRLTSPFTSGASQKILITNYGSSMIRLTNCDLSLHTFLAALFTHRFNLS  
CSEBRILASCR8IDHPDGMGPIITYAGSSSRPORYCMHVAAPRPGCIYPA5VGCPVYC  
PTSPVYVGTITDRGCVPTYSNGENETVILLNTRPPOQNMFGCTMANSNGFTTCGG  
PPCNIGAGANNLTCTCPDCEFRKPEEATYTKGSGSPWPLPRCMVYPTRLMHPCTVNP  
TIFRYRYVYGVERHLNAAACMTWGECCNEDRBSLSLPLLSSTWQJLIPCSFTPL  
PALSTGLIHLHQNVNVDQYLYGVSAVSVLYKWEYVLLFLFLADARVACIMMMLL  
IAQBALENIIVYNLAASVAGAGI1LSFLVFCAMVYIKGLVGAAYAFGWPVLL  
LILALPRAYAMDREMASSCGGCVGVGIVLITLSPYKVLFLAKLIMLYOYITBAEH  
LQWVPPLRNNGGDAILILTCANHPBLIPDITLLAILRGLMVLQAGITRVYFVR  
AQGLIRACMLVRKAYGGHYOMAFKLAALGTVYTHLTPLRMAHGLBLDAVAB  
PVFSDMEETK1ITWADTYACGDI1ISGIPVSRGRBILILPAPDSLGQQRHLLAPIT  
AYSQOTRGLGCIITSLTGRDXONVEBVOVASTRISFLATCVNGCMVYCHRAAGSK  
TLAPKPEIPTMYNVDODLVGMQAPGASRLTCTCGSSDLVYVTHADVYIPRRRG  
DSRGLSPRPVSYLKSGSSGPIILCP8GVGVIGIRAAVCTRGVAKAVDPEVAMEET  
MRSVPTDNSGPVAPOTPOVAHLAAITGSGKLVAAAYAAQCYKLVYINPSAATL  
GFGAYMSKATBDNIRTGIRITTTTGAPIITYSTYKTLADGGCGSGAYDIIICBCHS  
TDSPTIIGIIVLQDAETAGARLVVLTATPPGSAVYVPHNPIBEVALSNGETIPEYK  
AIPIEAIKGGRHILFCHSKKCCDELAKL8SLGINAAVAYKGLDVSYIPTSANVVA  
TDLMTGYTGDPSVIDNCTCVTVDPFSLDPTPTIETTVVPODAVGRTRGRG  
RGGIYRFVTPGERSGMFSDSVLCEYDAGCAMYELIPEATSVRLRAYLNTGDPVOC  
DHLEFMSVPTGLTHIDHPLSQTKGNDPPLYVAYOATYCARAQAAPPSMDMMFC  
LIRUKPLHGPPLYLIRGAVONSVTLTHPTTKYIMACMADELEVSTVYVGVYLA  
ALAAVCLTGSVIVIGRI1LSGRPAVLPDEBLVREFDEMEBCASHLYLEQMQLAB  
QFKOKALGLQATKQAEAAAPVESKMRALAEAFWAKMNFISGIQYLAGLSTLPGN  
PAISLMAFTASVTSPLTOSTLFLNIGGVAAQLAPPSAASAFVAGIAGAAGSI1  
GLKVLVDIILAGVAGVAGALVARKMSGWSPSTEDLVNLLPALSFGALVGVCAA  
ILRPHVPGAGAVOMNRILTAFA8RGHNVPTHYVPSDAARVTOILSLITPOLK  
RLHWINEDCSPTSCGSLMDVMDWICTVILIDFTWIOSKLPRLBVPFFSCORRYK  
GWRGDDIMHTTCCQAQIAGHVKN8MSR1VGPFTCSNTWMTGTPPINAVTGCTPSP  
APNYSRALMRYABEYEVTVRVGDHYITGQTDNINICPCQVAPABEPHETDGVRLR  
YAPACKPLIREAVFOYGLNQYLVG5OLPCGEPEDVAVLIT5MLTDPEHITAEYKRLR  
ARGSPDPMASASOLASPLSKAICTTH8SPDLLEANILAEOMEGNTRVRESN  
KVYILDSFDPDRAEDREVSVPARILKRN8PMPMPMPMPMPMPMPMPMPMPMPMPMP  
VPTVHGCPLPPTAAPPIPPRRKRTVYVLESTYSALAELATTESSBSASADCT  
ATAPDQASDDGDESVESYSMPPLEBEGDDLDLSDSVTSSEATBETVCCSWS  
YTWGALITPCAAESKLPINALNSK1LRHNMVATTSRSASLRQKVFEDRQVLD  
DHYVADLEKEMAKASYAKALSYEACKLPPTSPASAKFYGCAQVYANLSKAINH1  
R5VKQDLLEDTEPTIDITIMAKN8VFCQPKGGRKARLIVFDLGRVCEKMAVLD  
VSTGLROAVMSSYGRQYSPGRVPEYPLKTMKSSKCCQNGCRRRACRAGVLTSGNTL  
EBSIYQCCDIAPEARQI1RSLTEFLYIGGPTMNSKCCQNGCRRRACRAGVLTSGNTL  
TCYIKATPACCAALQDCTMLVNDDLVVLCESAGTQEDANLRFTEAMRTYSAPRG  
DPPOPEYDLELITSCSSNVSAHAD8AKRYVLLRDPITLTSRAAMTARTPNV5ML  
GNIIMYAPTLIARMIMTHF5IILAOOLEKALDQCOIYGACSI1EPLDQI1ERHL  
GLASPLSHSYSGEINRVA5CLRLGVPPLVWRH8RARSYAKLISQGRATQSKYL  
FMAVVRKTKLKTIPASOLDLSGWPFAVAGISGDIYHLS8ARR8MPLCILLVGV  
GIYILPNR"  
9293..9331  
/note="upstream of poly-U stretch"  
9332..9425  
/note="poly(U) length is indistinct  
poly(U) stretch"  
9426..9547  
/note="3' X tail"  
ORIGIN  
3'UTR  
9293..9331  
3'UTR  
9332..9425  
3'UTR  
9426..9547  
Beet Local Similarity 91.9%; Pred. No. 0;  
Matches 1663; Conservative 0; Mismatches 146; Indels 0; Gaps 0;  
QY 1 ATGAGCAGCAATCTCTAAACCTCAAGAAAACCAACGTAACACCAACCGCGCCACAG 60  
DB 260 ATGAGCAGCAATCTCTAAACCTCAAGAAAACCAACGTAACACCAACCGCGCCACAG 319  
QY 61 GACGTCAAGTTCGGGGCGGTGTGTCAGATCGTTGTGAGATTYACCTTGTCCGCGCAGG 120  
DB 320 GACGTCAAGTTCGGGGCGGTGTGTCAGATCGTTGTGAGATTYACCTTGTCCGCGCAGG 379  
QY 121 GAGCCCAAGTGGGNGTCGGCGCAGCTTGAAGAAAGCTTCGAGCGGTGCAACCTTGATGGA 180  
DB 380 GAGCCCAAGTGGGNGTCGGCGCAGCTTGAAGAAAGCTTCGAGCGGTGCAACCTTGATGGA 439  
QY 181 AGGCGACAACCTATCCCAAG3CTGCGCACGCCGAGAGGGCAGAGGCTCGGAGCTCAGCCCGGG 240

Db	440	AGGCGAACCTATCCCAAGGCTGCGAGCCGAGGGCAGAGGCTTGAGGCTTCAGCCGGG	499
Qy	241	TACCCCTTGGCCCTCTATATGCAACGAGGGGATGAGGTGGAGAGATGAGCTCTGTCAACC	300
Db	500	TATCCTTGGCCCTCTATATGCAATGAGGGCTGAGGGTGGGAGAGATGAGCTCTGTCAACC	559
Qy	301	GTGGCTCCGGGCTGATGTTGGGGCCCACTGACCCCGGGCGTAGGTCGCTGAATTTGGT	360
Db	560	CGTGGCTCCGGGCTGATGTTGGGGCCCACTGACCCCGGGCGTAGGTCGCTGAATTTGGT	619
Qy	361	AAAGTCATCGATACCCCTCACAATGGGGCTTGCCCACTCAATGGGGTAAATTCGCCCTGTC	420
Db	620	AAAGTCATCGATACCCCTCACAATGGGGCTTGCCCACTCAATGGGGTAAATTCGCCCTGTC	679
Qy	421	GGCCCTCCCTTAGAGGGGCGCTGCGCAGGGGCGCTGGGCGCATGAGCGTCCGGGTTCTGAGAGAC	480
Db	680	GGGCGCCCTTAGAGGGGCGCTGCGCAGGGGCGCTGGGCGCATGAGCGTCCGGGTTCTGAGAGAC	739
Qy	481	GGGCTGAACATATGCAACAGGAATTTACCCGAGTGTCTTTCTATCTTCTCTTGGCT	540
Db	740	GGGCTGAACATATGCAACAGGAATTTGCGCGGTTGTCTTTCTATCTTCTCTTGGCT	799
Qy	541	TTGCTGTCCTGTTTGAACATTCGAGCTCCGCTTATGAAAGCGCAACGTTCGGGATC	600
Db	800	TTGCTGTCCTGTTTGAACATTCGAGCTCCGCTTATGAAAGCGCAACGTTCGGGATC	859
Qy	601	TACCATGTCAGGAACGATTTGCTCCCACTCAAGCATCGTATCGAGAACAGCGGACATGATC	660
Db	860	TATCATGTCAGGAACGATTTGCTCCCACTCAAGCATTTGTATGAGCGGCGGACGTATC	919
Qy	661	ATGCAACACCCCGAGGTGTGTGTCCTGTGTCCGGAGGGTAATTCCTCCGCTGTGGGTA	720
Db	920	ATGCAATGCCCCCGAGGTGTGTGTCCTGTGTCCGGAGGGTAATTCCTCCGCTGTGGGTA	979
Qy	721	GCGCTCATCCCAAGCTCGCGGGCCCAAGAGCGCCAGCATCCCACTGGGACAATACGACGC	780
Db	980	GCGCTCATCCCAAGCTCGCGGGCCCAAGAGATTCACAGCGTCCCACTCAACAAATACGACGT	1039
Qy	781	CACGTCGATTTGCTCGTTGGGGGGGCTGCTTCGCTCGGCTATATGATGTAGGGAGATCTC	840
Db	1040	CACGTCGATTTGCTCGTTGGGGGGGCTGCTTCGCTCGGCTATATGATGTAGGGAGATCTC	1099
Qy	841	TGCGGATCTGTTTTCCTGCTCTGCTCAGCTGTTCACCTTCGCGCTCGCGCATACGACG	900
Db	1100	TGCGGATCTGTTTTCCTGCTCTGCTCAGCTGTTCACCTTCGCGCTCGCGCATACGACG	1159
Qy	901	GTAACGAGCTGCAATTTGTCATCTATTCGGGCGACGATATAGGTCACCGCATGGCTTGG	960
Db	1160	GTAACGAGCTGCAATTTGCTCACTATTCGGGCGACGATATAGGTCACCGCATGGCTTGG	1219
Qy	961	GATATGATGATGAAGTGGTCACTACACAGACCCAGTAGGTATGCGAGCTACTCCGATC	1020
Db	1220	GATATGATGATGAAGTGGTGGCCCAACAGACCCCTCGTGTGTGCACTACTCCGATC	1279
Qy	1021	CCACAAGCTGTCTGTGACATGAGTGGCGGGGCCCATGTGGGAGTCTTGGCGGGCTTGGC	1080
Db	1280	CCACAAGCTGTCTGTGACATGAGTGGCGGGGCCCATGTGGGAGTCTTGGCGGGCTTGGC	1339
Qy	1081	TACTATTTCCATGTGTGGGGAATGGGCTTAAGTCTTGGTGTGATCTCTTTGGCGGCG	1140
Db	1340	TACTATTTCCATGTGTGGGGAATGGGCTTAAGTGTGATGTACTCTTGGCGGCG	1399
Qy	1141	GTTACGGGGGAACCTTACACGACAGGGGGGACACAGCGGCGCGCGGACAGGGCTTACA	1200
Db	1400	GTTACGGGGGAACCTTACACGAGCGGGGGGCGGCAAGGCGCGAGCGCTTCCGECTCAGG	1459
Qy	1201	TCCCTTTTCAACCTTGGGCGGCTCAGAAATTCAGCTTGTATAACCAACGAGCGAGCTGG	1260
Db	1460	TCCCTTTTCAACCTTGGGCGGCTCAGAAATTCAGCTTATAAACAACATGAGCAGTTGG	1519
Qy	1261	CACATCAACAGATGCTTGAATGTGCAATGATCTCCCTCAAACTGGGTTCTTGGCGCG	1320

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL																																														
Db	AX739971	Sequence 2 from Patent WO02059321.	AX739971	AX739971.1	GI:30519245	Hepatitis C virus Hepatitis C virus Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.	de Francesco, R., Migliaccio, G. and Paolesse, G. Hepatitis C virus replicons and replicon enhanced cells Patent: WO 02059321-A 2 01-ANG-2002; ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGILETTI S. P. A. (IT)	1	ATGAGCAGCATCTTAACCTCAAGAAAAACAAACGTAACCAACACCGCCGACAG	60																																															
Db	342	ATGAGCAGCATCTTAACCTCAAGAAAAACAAACGTAACCAACACCGCCGACAG	401	61	GAGCTCAAGTTCCCGGGCGGTGTCAGATCGTGTGAGATTAACTGTGGCCGCGCAG	120																																																			
Qy	1570	CACATCAACAGGATCTGCCCTTAACCTGCATATACCTCCTCCACACTGGGTTCTTCCGCG	1579	1321	CTGTTCTTACACGACAGGTTCAATGCGTCGCGATGCTCAGAGCGCATGAGCGAGCTGCGC	1380	1580	CTGTTCTTACACACAGGTTCAACGCGTCGGATGTCGCGAGCGCCTTGAGCTGAGCTGCGC	1639	1381	CCCATTAACCAAGTTGATCAGAGGTTGGGTTCCCATCTACTTATTAATGATCCCAAGCTTG	1440	1640	TCATATGACAGGTTGATCAGAGGATGGGGTCCCATCACTTAATGCTCAACCTTACAGAGCTCG	1699	1441	GACCAAGAGGCGCTTATGCTGACCTAACGCACTCAACCGTGTGGATCTGTCGCGCGTTG	1500	1700	GACCAAGAGGCGCTTATGCTGACCTAACGCACTCAACCGTGTGGATCTGTCGCGCGTTG	1759	1501	CAGGTGTGTGAGCCAGTGTACTGTTTCACTTCAAGCCCTGTTTGTGTGGGACGACCGAT	1560	1760	CAGGTGTGTGATCCAGTGTATTTGTTTCAACCCCAAGCCCGGTTGTGTGGGACGACCGAT	1819	1561	CGTTTCGGCGCCCTTACGTAACATGTGGGTGAAATGACGACGTCGCTCTTCTCAAC	1620	1820	CGTTTCGGGTGTCCTTACGTAACATGTGGGTGAAATGACGACGTCGCTCTTCTCAAC	1879	1621	AACACGCGCGCCGACGCGGCGCACTGGTGTGGCTGTACATGAATGAATPAGACCGGGTTC	1680	1880	AACACGCGCGCCGACGCGGCGCACTGGTGTGGCTGTACATGAATGAATPAGACCTGGTTC	1939	1681	ACCAAGACGTGTGGGGGCCCCCGGTGACACATCGGGGGGCTCGGCAACACACTTTGATC	1740	1940	ACCAAGACGTGTGGGGGCCCCCGGTGACATCGGGGGGGCGGCAACACACTTTAACA	1999	1741	TGCCCCACGACGTGCTTCCGGAAGCATTCGAGGCCACTTACACCAATGCGGTTCCGGG	1800	2000	TGCCCCACGACGTGCTTCCGGAAGCACCCGACGACCTTACACCAATGCGGTTCCGGC	2059	1801	CCTTGGTTG	1809	2060	CCTTGGTTG	2068
Qy	1570	CACATCAACAGGATCTGCCCTTAACCTGCATATACCTCCTCCACACTGGGTTCTTCCGCG	1579	1321	CTGTTCTTACACGACAGGTTCAATGCGTCGCGATGCTCAGAGCGCATGAGCGAGCTGCGC	1380	1580	CTGTTCTTACACACAGGTTCAACGCGTCGGATGTCGCGAGCGCCTTGAGCTGAGCTGCGC	1639	1381	CCCATTAACCAAGTTGATCAGAGGTTGGGTTCCCATCTACTTATTAATGATCCCAAGCTTG	1440	1640	TCATATGACAGGTTGATCAGAGGATGGGGTCCCATCACTTAATGCTCAACCTTACAGAGCTCG	1699	1441	GACCAAGAGGCGCTTATGCTGACCTAACGCACTCAACCGTGTGGATCTGTCGCGCGTTG	1500	1700	GACCAAGAGGCGCTTATGCTGACCTAACGCACTCAACCGTGTGGATCTGTCGCGCGTTG	1759	1501	CAGGTGTGTGAGCCAGTGTACTGTTTCACTTCAAGCCCTGTTTGTGTGGGACGACCGAT	1560	1760	CAGGTGTGTGATCCAGTGTATTTGTTTCAACCCCAAGCCCGGTTGTGTGGGACGACCGAT	1819	1561	CGTTTCGGCGCCCTTACGTAACATGTGGGTGAAATGACGACGTCGCTCTTCTCAAC	1620	1820	CGTTTCGGGTGTCCTTACGTAACATGTGGGTGAAATGACGACGTCGCTCTTCTCAAC	1879	1621	AACACGCGCGCCGACGCGGCGCACTGGTGTGGCTGTACATGAATGAATPAGACCGGGTTC	1680	1880	AACACGCGCGCCGACGCGGCGCACTGGTGTGGCTGTACATGAATGAATPAGACCTGGTTC	1939	1681	ACCAAGACGTGTGGGGGCCCCCGGTGACACATCGGGGGGCTCGGCAACACACTTTGATC	1740	1940	ACCAAGACGTGTGGGGGCCCCCGGTGACATCGGGGGGGCGGCAACACACTTTAACA	1999	1741	TGCCCCACGACGTGCTTCCGGAAGCATTCGAGGCCACTTACACCAATGCGGTTCCGGG	1800	2000	TGCCCCACGACGTGCTTCCGGAAGCACCCGACGACCTTACACCAATGCGGTTCCGGC	2059	1801	CCTTGGTTG	1809	2060	CCTTGGTTG	2068
Qy	1570	CACATCAACAGGATCTGCCCTTAACCTGCATATACCTCCTCCACACTGGGTTCTTCCGCG	1579	1321	CTGTTCTTACACGACAGGTTCAATGCGTCGCGATGCTCAGAGCGCATGAGCGAGCTGCGC	1380	1580	CTGTTCTTACACACAGGTTCAACGCGTCGGATGTCGCGAGCGCCTTGAGCTGAGCTGCGC	1639	1381	CCCATTAACCAAGTTGATCAGAGGTTGGGTTCCCATCTACTTATTAATGATCCCAAGCTTG	1440	1640	TCATATGACAGGTTGATCAGAGGATGGGGTCCCATCACTTAATGCTCAACCTTACAGAGCTCG	1699	1441	GACCAAGAGGCGCTTATGCTGACCTAACGCACTCAACCGTGTGGATCTGTCGCGCGTTG	1500	1700	GACCAAGAGGCGCTTATGCTGACCTAACGCACTCAACCGTGTGGATCTGTCGCGCGTTG	1759	1501	CAGGTGTGTGAGCCAGTGTACTGTTTCACTTCAAGCCCTGTTTGTGTGGGACGACCGAT	1560	1760	CAGGTGTGTGATCCAGTGTATTTGTTTCAACCCCAAGCCCGGTTGTGTGGGACGACCGAT	1819	1561	CGTTTCGGCGCCCTTACGTAACATGTGGGTGAAATGACGACGTCGCTCTTCTCAAC	1620	1820	CGTTTCGGGTGTCCTTACGTAACATGTGGGTGAAATGACGACGTCGCTCTTCTCAAC	1879	1621	AACACGCGCGCCGACGCGGCGCACTGGTGTGGCTGTACATGAATGAATPAGACCGGGTTC	1680	1880	AACACGCGCGCCGACGCGGCGCACTGGTGTGGCTGTACATGAATGAATPAGACCTGGTTC	1939	1681	ACCAAGACGTGTGGGGGCCCCCGGTGACACATCGGGGGGCTCGGCAACACACTTTGATC	1740	1940	ACCAAGACGTGTGGGGGCCCCCGGTGACATCGGGGGGGCGGCAACACACTTTAACA	1999	1741	TGCCCCACGACGTGCTTCCGGAAGCATTCGAGGCCACTTACACCAATGCGGTTCCGGG	1800	2000	TGCCCCACGACGTGCTTCCGGAAGCACCCGACGACCTTACACCAATGCGGTTCCGGC	2059	1801	CCTTGGTTG	1809	2060	CCTTGGTTG	2068
Qy	1570	CACATCAACAGGATCTGCCCTTAACCTGCATATACCTCCTCCACACTGGGTTCTTCCGCG	1579	1321	CTGTTCTTACACGACAGGTTCAATGCGTCGCGATGCTCAGAGCGCATGAGCGAGCTGCGC	1380	1580	CTGTTCTTACACACAGGTTCAACGCGTCGGATGTCGCGAGCGCCTTGAGCTGAGCTGCGC	1639	1381	CCCATTAACCAAGTTGATCAGAGGTTGGGTTCCCATCTACTTATTAATGATCCCAAGCTTG	1440	1640	TCATATGACAGGTTGATCAGAGGATGGGGTCCCATCACTTAATGCTCAACCTTACAGAGCTCG	1699	1441	GACCAAGAGGCGCTTATGCTGACCTAACGCACTCAACCGTGTGGATCTGTCGCGCGTTG	1500	1700	GACCAAGAGGCGCTTATGCTGACCTAACGCACTCAACCGTGTGGATCTGTCGCGCGTTG	1759	1501	CAGGTGTGTGAGCCAGTGTACTGTTTCACTTCAAGCCCTGTTTGTGTGGGACGACCGAT	1560	1760	CAGGTGTGTGATCCAGTGTATTTGTTTCAACCCCAAGCCCGGTTGTGTGGGACGACCGAT	1819	1561	CGTTTCGGCGCCCTTACGTAACATGTGGGTGAAATGACGACGTCGCTCTTCTCAAC	1620	18																										



|||||  
Db 402 GAGCTCAAGTCCCGGCGGTGTGAGATCTCGGTGAGATTACCTGTTGCCGCGCAGG 461  
121 GGGCCCCAGGTGGTGTGGCGCGCATAGAGACTTCGAGCGGTCCGAACCTGTGTA 180  
Db 462 GGGCCCCAGGTGGTGTGGCGCGCATAGAGACTTCGAGCGGTCCGAACCTGTGTA 521  
181 AGGCGACAACCTATCCCGAAGGCTCGCCAGCGCGAGGAGGCGCTGGGCTCAGCCGG 240  
Db 522 AGGCGACAACCTATCCCGAAGGCTCGCCAGCGCGAGGAGGCGCTGGGCTCAGCCGG 581  
241 TACCTTGGGCGCTCTATGCGACAAGGCGATGGGGTGGGCAAGATGCTCTGTACCC 300  
Db 582 TACCTTGGGCGCTCTATGCGACAAGGCGATGGGGTGGGCAAGATGCTCTGTACCC 641  
301 CGTGGCTCCGGCTAGTGGGGCCCACTGACCCCGCGGTAGGTCGCTAATTGGGT 360  
Db 642 CGTGGCTCCGGCTAGTGGGGCCCACTGACCCCGCGGTAGGTCGCTAATTGGGT 701  
361 AAGTCATGATACCTCATAGCGGCTTCGCGCATTCATGGGGTACATTCGCTGTC 420  
Db 702 AAGTCATGATACCTCATAGCGGCTTCGCGCATTCATGGGGTACATTCGCTGTC 761  
421 GCGCTCCCTTAAAGGGGCGCTGCCAGAGGCTTCGGCGCATGCGCTCGGGTTCTGAGAG 480  
Db 762 GCGCTCCCTTAAAGGGGCGCTGCCAGAGGCTTCGGCGCATGCGCTCGGGTTCTGAGAG 821  
481 GGGGTGAATGACAAAGGGAATTTACCGGGTGTCTTCTATCTTCTCTGCTGCT 540  
Db 822 GGGGTGAATGACAAAGGGAATTTACCGGGTGTCTTCTATCTTCTCTGCTGCT 881  
541 TTGCTGCTCTGTTGACCAATCCAGCTTCGCTTATGAGTGGAGCGCAAGTGTCCGGAGTC 600  
Db 882 TTGCTGCTCTGTTGACCAATCCAGCTTCGCTTATGAGTGGAGCGCAAGTGTCCGGAGTC 941  
601 TACCATGTACGAAGATGCTTCCTCACTCAAGCATGCTGTACGAGACAGCGACATGATC 660  
Db 942 TACCATGTACGAAGATGCTTCCTCACTCAAGCATGCTGTACGAGACAGCGACATGATC 1001  
661 ATGCAACCCCGGGGTGTGCGCTGTCCGGGAGGATATCCCGCGCTGCTGGGTA 720  
Db 1002 ATGCAACCCCGGGGTGTGCGCTGTCCGGGAGGATATCCCGCGCTGCTGGGTA 1061  
721 GCGCTCACTCCACAGCTCGCGGCGCAAGAGCGCAGCATCCCACTGCGACAATAGACGC 780  
Db 1062 GCGCTCACTCCACAGCTCGCGGCGCAAGAGCGCAGCATCCCACTGCGACAATAGACGC 1121  
781 CACGTGATTTGCTGTTGGGCGGCTGCTTCTGCTCGCTATGATCGTGGGAGATCTC 840  
Db 1122 CACGTGATTTGCTGTTGGGCGGCTGCTTCTGCTCGCTATGATCGTGGGAGATCTC 1181  
841 TGGGATCTGTTTCTGCTCTCAAGCTTCACTTCCGCTCCGCGGAGATCAGAG 900  
Db 1182 TGGGATCTGTTTCTGCTCTCAAGCTTCACTTCCGCTCCGCGGAGATCAGAG 1241  
901 GTACAGAGCTGAATTTCAATCTATCCCGGCGAGTATCAGGTCAAGCGATGGCTGG 960  
Db 1242 GTACAGAGCTGAATTTCAATCTATCCCGGCGAGTATCAGGTCAAGCGATGGCTGG 1301  
961 GATATGATGATACTGTACCTTACAGAGCCCTAGTGTATTCGAGCTACTCCGATC 1020  
Db 1302 GATATGATGATACTGTACCTTACAGAGCCCTAGTGTATTCGAGCTACTCCGATC 1361  
1021 CCACAGAGCTGTGTGACATGTGTGGGGGGGCCACTGGGAGATCTTGCGGGGCTTGGC 1080  
Db 1362 CCACAGAGCTGTGTGACATGTGTGGGGGGGCCACTGGGAGATCTTGCGGGGCTTGGC 1421  
1081 TACTATTCATGATGGGGAAGTGGGCTAAGTCTTGTTGATGCTACTCTTTGGCGGC 1140  
Db 1422 TACTATTCATGATGGGGAAGTGGGCTAAGTCTTGTTGATGCTACTCTTTGGCGGC 1481  
1141 GTTGAAGGGAACCTTACAGAGAGGGGGAACACAGGCGCGCGCCACAGGGGCTTACA 1200  
|||||

Db 1482 GTTGAAGGGAACCTTATGTAGACAGGGGGAAGATGGCCAAAACACCCCTGGGATTAAG 1541  
1201 TCCCTCTTACACCTTGGGCGGCTCAGAAAATCCAGCTTGTAAACCAACGAGGAGCTGG 1260  
Db 1542 TCCCTCTTATACCGGGGTGATCCAGAAAATCCAGCTTGTAAACCAACGAGGAGCTGG 1601  
1261 CACATCAACAGAACTGCTTGAATGCAATGATCTCCCTCAAACTGGGTTCTTGGCGCG 1320  
Db 1602 CACATCAACAGAACTGCTTGAATGCAATGATCTCCCTCAAACTGGGTTCTTGGCGCG 1661  
1321 CTGTTCTACAGCAAGTTCATGCTCCGATGCTCAAGGCGCATGGCCAGCTGCGCG 1380  
Db 1662 CTGTTCTACAGCAAGTTCATGCTCCGATGCTCAAGGCGCATGGCCAGCTGCGCG 1721  
1381 CCCATTGACAGTTGATGATCAGAGGGGTGGGCTCCATCACTTAATGAGTCCAGGCTTG 1440  
Db 1722 CCCATTGACAGTTGATGATCAGAGGGGTGGGCTCCATCACTTAATGAGTCCAGGCTTG 1781  
1441 GACCAAGGCGCTTATGCTGGCACTACGACCTCAACCGGTGTATGTCGCCGCTTG 1500  
Db 1782 GACCAAGGCGCTTATGCTGGCACTACGACCGCGCGGTGTATGTCGCCGCGCG 1841  
1501 CAGGTGTGTGGCCCAAGTGTACTGTTTCACTCAAGCCCTGTTGTGTGGGAGACGAGAT 1560  
Db 1842 CAGGTGTGTGTGCAAGTGTACTGTTTCACTCAAGCCCTGTTGTGTGGGAGACGAGAT 1901  
1561 CGTTTGGGCGGCGCTTACAGATGAGAGGGGTGAGATGAGACGAGCGTGTCTTCAAC 1620  
Db 1902 CGTTTGGGCGGCGCTTACAGATGAGAGGGGTGAGATGAGACGAGCGTGTCTTCAAC 1961  
1621 AACAGCGGCGCGCAACGAGGCACTGTTGCGCTGTATGATGATGATGACGAGCGGTTTC 1680  
Db 1962 AACAGCGGCGCGCAACGAGGCACTGTTGCGCTGTATGATGATGATGACGAGCGGTTTC 2021  
1681 ACCAAGAGTGTGGGCGGCGCGCGCAACATCGGCGGAGGCTGGGCAACAACATTTGATC 1740  
Db 2022 ACCAAGAGTGTGGGCGGCGCGCGCGCAACATCGGCGGAGTGGCAATTAACCTTGAAC 2081  
1741 TGCCCCACAGAGCTGTTCCGAGAGATCCCGAGGCACTTACACCAATGCGGTTCCGGG 1800  
Db 2082 TGCCCCACAGAGCTGTTCCGAGAGATCCCGAGGCACTTACACCAATGCGGTTCCGGG 2141  
1801 CTTGGTTG 1809  
Db 2142 CTTGGTTG 2150

RESULT 13  
HCU238799  
LOCUS Hepatitis C virus type 1b complete genome, isolate Conl.  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
AU238799.1 GI:5420376  
complete genome; core protein; glycoprotein E1; glycoprotein E2;  
NS2 proteinase; NS3 proteinase/helicase; NS3/4A proteinase  
cofactor; NS4b protein; NS5A phosphoprotein; NS5B RNA dependent RNA  
polymerase; p7 peptide; polyprotein.  
SOURCE  
ORGANISM  
Hepatitis C virus type 1b  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepadnavirus.  
REFERENCE  
1  
AUTHORS Lohmann V., Korner F., Koch J., Heitman U., Theilmann L. and  
Bartenschlager R.  
TITLE Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell  
line  
JOURNAL Science 285 (5424), 110-113 (1999)  
MEDLINE 99322193  
PUBMED 10390360  
REFERENCES 2 (bases 1 to 9605)  
AUTHORS Bartenschlager R.  
TITLE Direct Submision  
JOURNAL Submitted (10-MAY-1999) Bartenschlager R., Institute for Virology,

Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67,  
55131 Mainz, GERMANY  
Location/Qualifiers  
1. .9605  
/organism="Hepatitis C virus type 1b"  
/viralion  
/mol\_type="genomic RNA"  
/strain="RB"  
/isolate="Con1"  
/db\_xref="taxon:31647"  
/country="Germany"  
1. .341  
/evidence=experimental  
342. .9374  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="CAB46677.1"  
/db\_xref="GI:5420377"  
/db\_xref="GOA:Q9MWX2"  
/db\_xref="UniProt/TREMBL:Q9MWX2"

5'UTR  
CDS

```
/translation="MSTNPKPKTKANTRRRQDYKFGGQIVGGVILPFRGRL
GVAATKTSERSQPRKQPIPAKQPEGAMAPQYPMPLYNBEGIGMAGWILSRG
SRPSMGPTDERRSRNIGKVIDTLGADLMGIVLGVAPLGAARALAGRVLED
GVNATGNLPGCSFSIFLLALSLCTIPASAYEVNVSQGVYHTVNDLCSNASIYEAD
MIMHTGCVPCVRENSSRCVVALPTLAARNASVPTTIRRHVDLLVGAALCSAMY
VGLDGSVFLVQLFTSPRRHETDQCSYRGHTRHAMMMMSPRALIV
SOLLRTPQAVYDMVNAHNGVLAGLAYISVNGMAVYLVMLLPAGVDGTYTGTM
AKNTLITLSPFGSSQKIQLVNTGSMHINRNLNCSNLGFLAALFVYKFNSS
GGERNASCSPIDAFAQGWGPIITYNESHSDORPYOMHAPRCQIVPAQVGPYC
FTSPVAVGTDRDFGVPTYSNGENETDVLNTRPPQGNMFCGTWNSTGFTKCG
PCNIGIGKTKLCPDCEKREKAEATYKSGPMPLTRCLVHYPRYLWYCTVNF
TTEKVMYGVGVLEHLEACNMRGERCNLEDBRELSPLSTTEMVYRCSFTTL
PALSTGLIHQNVVDVYIYIGSAVSFAIKWEVYLLFLLADARCACLMMHLL
LQAEALLENLVNNAASVAGAGHILSFVFCAAVYIKGRVPGAVALYGVPLL
LIALPRAVYANDREMAASCGAVFVGLILTLSPHYKLFARLIMWLYFTIRAH
LQWIPPLNVGRGRDAVILITCAIHBEILFTIKILAILGMLVLAQITKVPYR
AGLIRACMLVRKAGVYOMALALITGVYVDHLTPLRDMAHAGRLDAVE
PVFESDMETKVIITGDRNQAACDITLGLPVSARGREIHLGPADSLGGCMRLAIRT
AYSQTRGLGCTITSLTGRDRNQGEVQVYSTANOSFLATVNGVCTVHGASR
TLAEPKPIITQMTNVDDLVGMQAPPGASLTPCTGSSDLVLTVRHADVLEVRG
DSRGLSPRPVSYLFGSSGGLCPSGHAGVIFRAVCTRGAKAVDVPVSMETT
MRSPTDNDSPPAVPOTFOVAHLHAPTSKSTKPAVAAGYKVLVNSVATL
GFCAYMSKAHGDIPNIRGTGRTITGAPITYSYGEKLDGGCGGAYOIIICDECHS
TOSTTIGTVLDQAEAGALVYATATPSPSVVPHNTEVSLSTGERPYGK
AIPETIKGRHLIFCHSKKCDLAKISGLCLNAVAYRGLDVSITSGDIVA
TDLMTGFTGDESDVIDCNVYQVDFSLDPTLETTLVTPQDAVRSQRRTG
RMGIYRFTVPERSGMDSVLCCEYDAGCANFPLVAYQATVCARQAQAPPMDQMC
DLHEFESVFTGLIHDAHFLSOTKQAGNFPYLVAYQATVCARQAQAPPMDQMC
LRLKPTLHGPTPLRGLAVONEVTTHTPIKYIMACMSADLEVVTSTWLVGYLA
ALAAVCLTGSVVIYGRITLSGRPAIIPREVIYREFDEMEBCASHLPIBGMQALAE
QKOKAIGLQTNKQEAAPVRESKMTLEFMAKHMNFIISGIOYLAGSLTEGN
PAIASLMAFTASTISPLTQHTLLENIIGGWAQAQAPSAASAFAGAIAGSLTEGN
IGKXVLVDILAGAGVAGALVAFKVMSEMSSTEDLVNLLPILSPGALVGVCA
ILIRHVPEGAGVOMNRLIAFASRGNHSPHYVESPAARVQIISLITITOLK
RHOWINENCTPCGCAQITGHYKNSMAIVGPTGNTWHTGFPFNATTTGCTPS
GVMRGDGMQTTCCPGCAQITGHYKNSMAIVGPTGNTWHTGFPFNATTTGCTPS
ANYSBALMRVAABEYEVTRGVDFVYVGMDDNVKCCQVPAPEFTEVGVRLHR
YAPACKPILREVEITLVGLNOYLVSGLCEEPDVAVLTMLTDSHITATKARL
AGSPSPILASSASQSLASPLKATCTTRHSDPADILEANILMRQSGNITRVEEN
KVIILDSFEPLQAEDEBEREVSVAELIIRSKRPRAMPINARDVNPPLSEMGDQY
VPPVHGCPLPPAKAPIIPPRKRTVYLSSEKVSALALAKTKTPGSESSANDGT
ATASPDQSDDDAGSDVESYSMPLEBEPDDPLSDSKSTVSEBESAEVYCCMS
YTWGLITPCAEETKLPINALSNSLHLHNHLYATVTRSLARKKVTARLOVLD
DHRYDLKEMKAKASTVAKLSEBAKCLTPHSAKSFEGYAGKVRNLSKRAVLI
RSWKLLEDETETPIDITIMARKNEVCVPEKGRKDALIYEPDGVNCEKMALYD
VSTLQAVMGSSSYGFOYSPQGRVEFLVNAKAKKCPAFADITPDFTVENDIRV
BESIYOCCLAPBARQATIRSLTERLYIGPLNSKGNCGNRCARCSGLTYSNCGTL
TCYKAAACRAKAKODCTMLVCGDLVYICESAGQDEBALSIRATEMTKYSAPG
DPRKPYDLELTSCSNVSVAHDSGRVYVUTRDPITPLAARAEKARHPVNSWL
GNINAPFTLMARMLIMTHFISILAOELERALKDCQIAGVACSILEPDLPIITRLH
GISAPLHSYSGEINRVAISCRLKGVPLRWRHARBARVRLSOGERRALTCGYL
FMAVATTKLKTPIPAASQDLISWFWAGYSGDIIYHSIRARPRFWMCLLLISGV
GYILPNR"
```

	mat_peptide	342. .914 /gene="core" /product="core protein"
	mat_peptide	915. .1490 /gene="E1" /product="glycoprotein E1"
	mat_peptide	1491. .2579 /gene="E2" /product="glycoprotein E2"
	mat_peptide	2580. .2768 /gene="p7" /product="p7 protein"
	mat_peptide	2769. .3419 /gene="NS2" /product="NS2 proteinase"
	mat_peptide	3420. .5312 /gene="NS3" /product="NS3 proteinase/helicase"
	mat_peptide	5313. .5474 /gene="NS4A" /product="NS3/4A proteinase cofactor"
	mat_peptide	5475. .6257 /gene="NS4B" /product="NS4B protein"
	mat_peptide	6258. .7598 /gene="NS5A" /product="NS5A phosphoprotein"
	mat_peptide	7599. .9371 /gene="NS5B" /product="NS5B RNA dependant RNA polymerase"
	gene	342. .914 /gene="core" /gene="E1" /gene="E2" /gene="p7" /gene="NS2" /gene="NS3" /gene="NS4A" /gene="NS4B" /gene="NS5A" /gene="NS5B" /evidence=experimental
ORIGIN	3'UTR	9374. .9605
Query Match	86.9%; Score 1572.2; DB 14; Length 9605;	
Best Local Similarity	91.8%; Fred. No. 0;	
Matches 1661; Conservative	0; Mismatches 148; Indels 0; Gaps 0;	
QY	1 ATGACGACGAATCTTAACCTCAAGAAACCAACGTAACCAACCGCGCCGACG	60
DB	342 ATGACGACGAATCTTAACCTCAAGAAACCAACGTAACCAACCGCGCCGACG	401
QY	61 GACGTCAAGTCCCGGCGGTGTCAGATCGTGTGAGATTAACTGTGCGCGCAG	120
DB	402 GACGTCAAGTCCCGGCGGTGTCAGATCGTGTGAGATTAACTGTGCGCGCAG	461
QY	121 GGGCCCAAGTGGGTGCGCGCGCACTAGAGAACTTCCGACGGTCCCAACTGTG	180
DB	462 GGGCCCAAGTGGGTGCGCGCGCACTAGAGAACTTCCGACGGTCCCAACTGTG	521
QY	181 AGCGCAACAACCTATCCCAAGGCTGCGCAGCCGAGGCGCAAGGCTGAGCCGCG	240
DB	522 AGCGCAACAACCTATCCCAAGGCTGCGCAGCCGAGGCTGAGGCTGAGCCGCG	581

OY	241	TACCTTGGCCCCCTGATAGGACAAGAGGCAATGGAGTGGGAGAGATAGGCTCGTCACCC	300
Db	582	TACCCCTGGCCCCCTTATAGCAATAGAGGCTTGGAGTGGGAGATAGGCTCTGTGCACCC	641
OY	301	CGTGGCTCCCGGCTAGTTGGGGCCCACTGACCCCGGCGATAGGTGCGTAAATTTGGAGT	360
Db	642	CGTGGCTCTCGGCTAGTTGGGGGCCCAAGGACCCCGGCGTAAAGTGGCGCAATTTGGAGT	701
OY	361	AAAGTCATCGAATACCCCTCAATAGGGGCTTGGCGACATCATTGGGGTATCAATTCGGCTCGTC	420
Db	702	AAAGTCATCGAATACCCCTCAAGCGGGCTTGGCGCATCTCATGAGGATCAATTCGGCTCGTC	761
OY	421	GGCGCTCCCTTATGGGGGCGCTGCACAGGGCCCTGGGCGCATAGCGTCCGGATTTGAGAGAC	480
Db	762	GGCGCCCCCTTAGGGGGCGCTGCAGGGGCCCTGGGCGCATAGCGTCCGGATTTGAGAGAC	821
OY	481	GGCGTGAACATATGCACAAGAGGAATTTACCGGTTGCTCTTCTATCTTCTCTTGGCT	540
Db	822	GGCGTGAACATATGCACAAGGGGAATCTGCCGGTTGCTCTTCTTATCTTCTCTTTGGGCT	881
OY	541	TTCGTGTCCTGTTTATGACCATTCAGGTTCCGGTTATGAAGTGGCAAGTTCGGAGATC	600
Db	882	TTCGTGTCCTGTTTATGACCATTCAGGTTCCGGTTATGAAGTGGCAAGTATCCGGAGT	941
OY	601	TACCATGTCAAGAACGATTTGCTCAACTCAGACATCTGTATGAGAGACAGCGGACATGATC	660
Db	942	TACCATGTCAAGAACGATGCTCCAACAGACATCTGTATGAGAGCGGACATGATC	1001
OY	661	ATGCACACCCCCGGGATGTGCCCCGTGTGTCGGGAGGGTAAATTCCTCCGCGTGGGTA	720
Db	1002	ATGCACACCCCCGGGATGTGCCCCGTGTGTCGGGAGGAACAATCTCCTCCGCTGGGTA	1061
OY	721	GCGCTACACTCCACGCTCGCGGCGCAAGGACGCGCAGATCCCACTGCGCAATTAAGACGCG	780
Db	1062	GCGCTACACTCCACGCTCGCGGCGCAAGGACGCTAGGATGCCCACTAAGCAGATTAAGACGCG	1121
OY	781	CACGTCGATTTGCTGTTGGGGCGGCTGCTTGTGCTCCGTAATGTAAGTGGGAGATCTC	840
Db	1122	CATGTCGATTTGCTGTTGGGGCGGCTGCTTGTGCTCCGTAATGTAAGTGGGAGATCTC	1181
OY	841	TGCGGATCGTTTCTCGTCTCAGCGTTCACCTTCGCGTCGCGCAATCAGACG	900
Db	1182	TGCGGATCGTTTCTCGTCTCGCGCCAGCGTTCACCTTTCGCGTCGCGGCAAGACG	1241
OY	901	GTACAGACTGCATTTGTTCAATCTATCCCGGCGACGATCAAGTCAACCGCATGGCTTG	960
Db	1242	GTACAGACTGCATTTGTTCAATATATCCCGGCGACGATCAAGTCAACCGTATGGCTTG	1301
OY	961	GATATGATGATGAACCTGTCACCTTACAGAGGCCATAGTGTGCGAGCTACTCCGGATC	1021
Db	1302	GATATGATGATGAACCTGTCACCTTACAGAGGCCATAGTGTGCGAGTACTCCGGATC	1361
OY	1021	CCACAAGCTGTGTGGACATGTGTGGCGGGGCCCATGSGGAGTCTCTGGCGGACTTGGCC	1080
Db	1362	CCACAAGCTGTGTGGATATGTGTGGCGGGGCCCATTGGGGAGTCTCTAAGCGGCGCTGGC	1421
OY	1081	TACTATTCATAGTGGGGAACTGGGCTAAAGTCTTGTGTGATGTCTCTTTGCGGCG	1141
Db	1422	TACTATTCATAGTGGGGAACTGGGCTAAAGTCTTGTGATGTCTCTTTTGGCGGCG	1481
OY	1141	GTTGACGGGGAACTTACACGACAGGGGGGACACAGCGCGCGGCCCAAGGGGCTTAC	1201
Db	1482	GTTGACGGGGGAACTTATGTGACAGGGGGGACATAGGCCAAAAACACCTTCGGGATTAGC	1541
OY	1201	TCCCTCTTTCACACTTGGGCGGCTCAGAAATTCAGCTTGTAAACACAAAGGCGACGTGG	1261
Db	1542	TCCCTCTTTCACACCGGCGTATCCAGAAATTCAGCTTGTAAACACAAAGGCGACGTGG	1601
OY	1261	CACATCAACAGAACTGCTTGAATCGCAATGATCCCTCAACATGGGTTCTTGGCGGCG	1321
Db	1602	CACATCAACAGAACTGCTTGAATCTGAATGATCTCCTCAACATCGGATTCCTGTGTGG	1661

QY	1321	TTGTTTCAACACGACAGAGTTCAATAGCGTCCGGATGAGCTCAAGAGCGCATGGCCAGCTGCACGC	1380
Db	1662	CTGTGTTTCACTGACCAAGTTCTCACTCATCTGGAATGCCAGAGCCCAATGACCAGCTTGACAC	1721
QY	1381	CCCATTTGACCAAGTTTCGATTCAGAGGGGTGGGGGATCCATCACTTAATATGATGCCACAGGCTTG	1440
Db	1722	CCCATTCGACGGGTTTCGCTCAGAGGGTGGGGGCCCATCACTTAACATGATGATCAACAGGCTGC	1781
QY	1441	GACCAGAGGCCCTATTGTCTGGCACTACGCACTTCACCGTGTGATTCGTGCCCGCTTG	1500
Db	1782	GACCAGAGGCCCTTATTGTGTGGCACTACGCAACCCGGCCGTCGCTATTCGTATCCGCGACG	1841
QY	1501	CAGGTGTGTGGCCCAAGTGTACTGTTTCACTTCAGACCCCTGTTGTGTGTGGGACGACCCGAT	1560
Db	1842	CAGGTGTGTGTTCAGTGTACTGTCTTACCCCAAGCCCTGTCTGTGTGGGACGACCCGAC	1901
QY	1561	CGTTTCGGCGCCCTTACGTACAGATGGGGGTGAGATGAGACGGACGTGCTTCTCAAC	1620
Db	1902	CGGTTTCGGCGTCCCTACGTACAGTTGGGGGAGATGAGAGGACGTGTCTTCTTAAAC	1961
QY	1621	AACACGCGGCGCCGACGCGGGCAACTGTGTTGCGCTGTACATGATGAATAGCACCGGGTTC	1680
Db	1962	AACACGCGGCGCCGCGCAAGGACACTGTGTTTGGCTGTACATGATGAATAGCACTGGGTTTC	2021
QY	1681	ACCAAGACGTGTGGGGGCGCCCCGTGTCAACATCGGGGGGGTCGCGCAACACACTTTGATC	1740
Db	2022	ACCAAGACGTGTGGGGGCGCCCCGTGTAACTATCGGGGGGATTCGGCAATATAAACCTTGCAC	2081
QY	1741	TGCCCCACGACCTGCTTCCGGAAGCAATCCGAGGCCATTACACCAATATGGGTTTCGGGG	1800
Db	2082	TGCCCCACGGACTGCTTCCCGGAACACCCCGAGGCCACTTACACCAAGTGTGTTTCGGGG	2141
QY	1801	CCTTGCTTG 1809	
Db	2142	CCTTGCTTG 2150	

```

FEATURES
    LOCUS      AB008441          3296 bp      RNA      linear      VRL 04-SEP-1998
    DEFINITION Hepatitis C virus RNA for polyprotein, partial cds, clone:J5-2.
    ACCESSION  AB008441
    VERSION    AB008441.1  GI:3523046
    KEYWORDS   polyprotein.
    SOURCE      Hepatitis C virus
    ORGANISM    Hepatitis C virus
    VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
    Hepacivirus.
    1 (sites)
    Kato,N., Ikeda,M., Sugiyama,K., Mizutani,T., Tanaka,T. and
    Shimotohno,K.
    Hepatitis C virus population dynamics in human lymphocytes and
    hepatocytes infected in vitro
    J. Gen. Virol. 79 (Pt 8), 1859-1869 (1998)
    MEDLINE    98378035
    JOURNAL    9714233
    PUBMED     2 (bases 1 to 3296)
    REFERENCE  Kato,N.
    DIRECT SUBMISSION
    Submitted (18-OCT-1997) Nobuyuki Kato, National Cancer Center
    Research Institute, Virology Division; 5-1-1 Tsukiji, Chuo-ku,
    Tokyo 104, Japan (E-mail:nkato@nc.ri.go.jp,
    Tel:03-3542-2511 (ex.4701), Fax:03-3543-2181)
    LOCATION/Qualifiers

```

CDS

```

/codon_start=1
/product="polypeptidein"
/protein_id="BAA32647.1"
/db_xref="GI:3523047"
/translation="MSTNPKPQKTKRNTNRPRQDYKFPGGQIVGVYLLPRGRRL
GVRATKTERSQPRGRQPIPKARRPEGAAMAPQYEMPLYGNEBELAGWMLISRG
SRPSMGPTDPRRRSRNLGKVIDTLTCGFLDMGYILVAPALGGAPALAHGVRVLE
SNVATGNLPGCSFSLFILALISCLTI PASAYEVNVSQVYHVTNCSNLSIYEAD
MNMHPGCVPCVREANSRQWALPTLAAHNSVPTTTIRHVDLVLGAALCSAMY
VQDLGQSVLISQLPRESRPOYETVODCKSLYRPHSVIRHMMMDMMSPPLALVY
SOLLRIPQAVDMVAGHWGLAGLAYSINVGMAVYLVMILFACVDAATTSGSTV
GRTTGLTSFAPBPQKLIQINTNSMHNRTALNCNDSLHTGLAALFYHKRNS
GSEBMAICRPIDKFAQGWPIITGVDPDDPDRPCMYHAPRCGLVAPQYQVCPYC
FTSPVVGTTDRSGAPTYTNGENETDVLNNTRPQSNMFGCTVNSNLTGFTKCG
PCNIGVGNNTLTCPTDCPRKEATYTKCSGPMILTRCWMDYRILMHYPCYTF
TIFKRYMYGVGEHRLDAA CNMTRGRCHVEDRSELSPLSTTEMVULPQSTPL
PALSTGLIHLHONTIDVOYLTVGVSVVSVKMEYVLLFLLIADARCAICLMMMLL
VQABALENLVLVINAASVAGAGILSFVFECAAMYIKGLVPGAAPAYGVWPLLL
LILALPRAVYAMDREMAASCGGAVFVGLALLTSPHYKFLARLIMWLYLITRAAL
LQWMIPELVNVRGDRALILLTCVHPELIFDIKILALIEGPLMVLQAGLTRVPEVR
AGLIRVCMILVRKVGHYVQAMLMBLALITGVYVDHLTPLRGMAHAGRLDAVAVE
PVFSDMETKIITWGDYACGDIIIGLFPVSARRR"

```

## ORIGIN

Query Match 86.8%; Score 1570.6; DB 14; Length 3296;  
 Best Local Similarity 91.8%; Pred. No. 0;  
 Matches 1660; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

```

QY 1 ATGAGCAGAAATCTAAACCTCAAGAAAAAACAAGTAACCAACGCGCGCCACAG 60
DB 273 ATGAGCAGAAATCTAAACCTCAAGAAAAAACAAGTAACCAACGCGCGCCACAG 332
QY 61 GACGTAAATGCCCGGCGGTGTGTCAGATCGTTGGTGGAGTTTAACTGTTGCCGCGCAG 120
DB 333 GACGTAAATGCCCGGCGGTGTGTCAGATCGTTGGTGGAGTTTAACTGTTGCCGCGCAG 392
QY 121 GGCCTCAGGTGGGTGTGCGCGGAGTAAGAACTTCCGAGCGGTGCAACTCTGTGA 180
DB 393 GGCCTCAGGTGGGTGTGCGCGGAGTAAGAACTTCCGAGCGGTGCAACTCTGTGA 452
QY 181 AGGCGACAACCTATCCCAAGGCTGCGCAAGCCCGAGGGGCGGCGTGGGCTCAAGCCCGAG 240
DB 453 AGGCGACAACCTATCCCAAGGCTGCGCAAGCCCGAGGGGCGGCGTGGGCTCAAGCCCGAG 512
QY 241 TACCTTGGCCCTCTATGGCAACGAGGGATGGGGTGGGCGAGATGGCTCTGTCAACC 300
DB 513 TACCTTGGCCCTCTATGGCAATGAGGGCTTGGGGTGGGCGAGATGGCTCTGTCAACC 572
QY 301 CGTGGCTCCGCGCTAGTGGGCGCCCACTGACCCCGCGGCTAGTGGCTAATTTGGGT 360
DB 573 CGTGGCTCCGCGCTAGTGGGCGCCCACTGACCCCGCGGCTAGTGGCTAATTTGGGT 632
QY 361 AAAGTCATCGATACCTCTCAATGCGGCTTGCAGACTTCAATGGGGTAACTTTCGCTGTC 420
DB 633 AAAGTCATCGATACCTCTCAATGCGGCTTGCAGACTTCAATGGGGTAACTTTCGCTGTC 692
QY 421 GCGCGCTCCCTTAGGGGGCGTGCAGAGGCTTGGCGGCAATGGCGTCCGGGTTTGGAGAGAC 480
DB 693 GCGCGCTCCCTTAGGGGGCGTGCAGAGGCTTGGCGGCAATGGCGTCCGGGTTTGGAGAGAC 540
QY 481 GCGGTGAACATAGCAACAGGGAATTTAACCGGTTGCTCTTCTATCTTCTCTTGGGCT 540
DB 753 GCGGTGAACATAGCAACAGGGAATTTAACCGGTTGCTCTTCTATCTTCTCTTGGGCT 812
QY 541 TTGCTGTCTGTTTGAACATTCAGGCTTCCGTTATGAAATGGCGCAACGTTGCCGGAGTC 600
DB 813 TTGCTGTCTGTTTGAACATTCAGGCTTCCGTTATGAAATGGCGCAACGTTGCCGGAGTC 872
QY 601 TACCATGTACGAAAGATTTGCTCAACTCAAGCATGTGTACGAGACAGGAGCAATGATC 660
DB 873 TACCATGTACGAAAGATTTGCTCAACTCAAGCATGTGTACGAGACAGGAGCAATGATC 932
QY 661 ATGCACAACCCCGGGTGTGTGCTGTGTCCGGAGGGTAAATTCCTCCGCTGTGGGTAA 720

```

```

DB 933 ATGCACAACCCCGGGGTGCTTACCTGCTGCTCCGGAGGCCCAACTCTCCGCTGTGGGTAA 992
QY 721 GGGCTACCTCCCAAGCTGTGGGGCGCAAGAGCGGACAGATCCCACTGCGCAATATGACAGCG 780
DB 993 GGGCTACCTCCCAAGCTGTGGGGCGCAAGAGCGGACAGATCCCACTGCGCAATATGACAGCG 1052
QY 781 CAGTCGATTTTCTCGTTGGGGCGGCTGCTTGTGCTCCGCTATGATGACGTGGGGATCTTC 840
DB 1053 CAGTCGATTTTCTCGTTGGGGCGGCTGCTTGTGCTCCGCTATGATGACGTGGGGATCTTC 1112
QY 841 TGGGATCTGTTTCTCTGCTCTCTCAAGCTGTTCACTTTCGCTGCGCGCAATCAAGC 900
DB 1113 TGGGATCTGTTTCTCTGCTCTCTCAAGCTGTTCACTTTCGCTGCGCGCAATCAAGC 1172
QY 901 GTAACAGATGCAATTTGTTCAATCTATCCCGGCGACGTATCAAGTCAACGCAATGCTTGG 960
DB 1173 GTAACAGATGCAATTTGTTCAATCTATCCCGGCGACGTATCAAGTCAACGCAATGCTTGG 1232
QY 961 GATATGATGATGAATGTCACCTACAGAGCCCTAGTGGTATGCGAGCTACTCCGAGTC 1020
DB 1233 GATATGATGATGAATGTCACCTACAGAGCCCTAGTGGTATGCGAGTGTCTCCGAGTC 1292
QY 1021 CCACAAGCTGTGTGACATGTGTGGGGGGCCCACTGGGAGATCTGGCGGCTTGGCT 1080
DB 1293 CCACAAGCTGTGTGACATGTGTGGGGGGCCCACTGGGAGATCTTACGGGGCTTGGCT 1352
QY 1081 TACTATTCATGTGTGGGAACTGGGCTAAGTCTTGTGTGTGATGCTACTCTTGGCGGC 1140
DB 1353 TACTATTCATGTGTGGGAACTGGGCTAAGTCTTGTGTGTGATGCTACTCTTGGCGGC 1412
QY 1141 GTTGAAGGGGAACCTTACAGAGACAGAGGGGAGACACAGCGCGCGCCACAGGGCTTAA 1200
DB 1413 GTTGAAGGGGACCTTACAGAGACAGAGGGGAGACAGCGCGCGCCACAGGGCTTAA 1472
QY 1201 TCCCTCTTCAACCTTGGGGCGGCTCGAATAATTCACCTTGTAAACCAACGCGACGCTGG 1260
DB 1473 TCCCTCTTCAACCTTGGGGCGGCTCGAATAATTCACCTTGTAAACCAACGCGACGCTGG 1532
QY 1261 CACATTAACAGAACTGCTTGAATCTGCAATGACTCCCTCCAACTGGGGTTCCTTGGCGGG 1320
DB 1533 CACATTAACAGAACTGCTTGAATCTGCAATGACTCCCTCCAACTGGGGTTCCTTGGCGGG 1592
QY 1321 CTGTTTACAGCAGAGGTTTCAATGCGTCCGAGTGTCTGAGGCGATGGCCAGCTCCGC 1380
DB 1593 CTGTTTACAGCAGAGGTTTCAATGCGTCCGAGTGTCTGAGGCGATGGCCAGCTCCGC 1652
QY 1381 CCCATTGACCAAGTTCATCAGAGGGGTGGGGTCCCATCACTTAATATGATCCCAAGCTTGG 1440
DB 1653 CCCATTGACCAAGTTCATCAGAGGGGTGGGGTCCCATCACTTAATATGATCCCAAGCTTGG 1712
QY 1441 GACAGAGGCTCTATTTGCTGGCACTACGCACTTCAACCGTGTGTATGTGTCGCCGCTGG 1500
DB 1713 GACAGAGGCTCTATTTGCTGGCACTACGCACTTCAACCGTGTGTATGTGTCGCCGCTGG 1772
QY 1501 CAGGTGTGTGGGCGCAGTGTACTGTTTCACTTCAAGCCGTTGTGTGGGGAGACAGCAT 1560
DB 1773 CAGGTGTGTGTGGGCGCAGTGTACTGTTTCACTTCAAGCCGTTGTGTGTGGGGAGACAGCAT 1832
QY 1561 CGTTTGGCGCCCTTACATAGATGGGGTGAAGATGAGACGAGCTGTGCTTCTTCAAC 1620
DB 1833 CGTTTGGCGCCCTTACATAGATGGGGTGAAGATGAGACGAGCTGTGCTTCTTCAAC 1892
QY 1621 AACACGGCGCCCGCACGGGGCAACTGTGTGCTGTATGATGATGATATGACACGGGCTTC 1680
DB 1893 AACACGGCGCCCGCACGGGGCAACTGTGTGCTGTATGATGATGATATGACACGGGCTTC 1952
QY 1681 ACCAAGAGTGTGGGGGGCGCCCGTGCACACATCGGGGGGTGGGCAACAACCTTGTATC 1740
DB 1953 ACCAAGAGTGTGGGGGGCGCCCGTGCACACATTTGGAGGGGTGGGCAACAACCTTGTATC 2012
QY 1741 TGGCCACAGGACTGCTTCCGGAAGCATCCGAGGCCACTTAAACCAATATCGGGTTGGGG 1800

```

Db 2013 TGCCCCAGGACTGCTTCGGAAGACCCCGAGGCCACTTACCAATCGGCTCGGG 2072  
Oy 1801 CCTGGTTG 1809  
|||  
|||  
Db 2073 CCATGTTG 2081  
  
RESULT 15  
LOCUS HCU238800  
DEFINITION Hepatitis C virus type 1b complete genome, isolate NCI.  
ACCESSION AJ238800.1 GI:5748510  
VERSION complete genome; core protein; glycoprotein E1; glycoprotein E2;  
KEYWORDS NS2 protease; NS3 proteinase/helicase; NS3/4A proteinase  
cofactor; NS4b protein; NS5A phosphoprotein; NS5B RNA dependant RNA  
polymerase; p7 peptide; polyprotein.  
SOURCE Hepatitis C virus type 1b  
ORGANISM Hepatitis C virus type 1b  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.  
REFERENCE 1  
AUTHORS Koch, J.O. and Bartenschlager, R.  
TITLE Modulation of hepatitis C virus NS5A hyperphosphorylation by  
nonstructural proteins NS3, NS4A, and NS4B  
J. Virol. 73 (9), 7138-7146 (1999)  
JOURNAL 99370154  
MEDLINE 10438800  
PUBMED  
REFERENCE 2 (bases 1 to 9033)  
AUTHORS Bartenschlager, R.  
TITLE Direct Submission  
JOURNAL Submitted (10-MAY-1999) Bartenschlager R., Institute for Virology,  
Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67,  
55131 Mainz, GERMANY  
FEATURES  
source location/Qualifiers  
1..9033  
/organism="Hepatitis C virus type 1b"  
/viroion  
/mol\_type="genomic RNA"  
/strain="R8"  
/isolate="NCI"  
/db\_xref="taxon:31647"  
/country="Germany"  
1..9033  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="CAB53095.1"  
/db\_xref="GI:5748511"  
/db\_xref="GOA:Q9Q066"  
/translation="MSTNPKPQRTKRTNRR.POVNKPFGGGQIVGVYL.PRRPRL  
GYRATKTSERSGRGRPIPKAQPRGRMAQGYWPLYNGEGIMAGMULSPRG  
SRPSGPTDPRKSRNLGKVIDLTTCGFDLMGYLPLGAPLGGARLHAGVRLD  
GNVYATGNLPGCSFSIPLIALISCLTIPASAVEVNVGAVHVNDCSNASIVYEAAD  
MIMTPGCGVPCVRENNSRCWALPTLAAANASPTTIRRHVDLILGAALCSAM  
VGDLCGVFLVAQLFTFSPRHEVTDNCISYPGHATGHRAMDMNMNPTALIVY  
SOLIRIPOAVUMVAAHMGVLAGLAVYMGVMAKVLIVMLPAGVDGGRVYCGTM  
AKNTIGITSLSPSGOKIQLVNTNGSHINRTALNCDSLNTGLALPVIHKNSS  
GCEPBMASCSPIDAQCGMGPITTVNESHSDRPYCMHYPAPCGIVPAADCGGVYC  
FTPSVVVGTDRFGVPTPYSGWENETDVLNNTSPGDMNGTANSTGTCKCG  
PPCNIGIGISKTLCTPCDFRGPRPATYTKCGSGPMLIPCLVHYPRIMHPCTVNF  
TIFKRVTVGGVHEHLEAAQWTRGERCNLEDRDSEILSLISTTEMOVLPCSPFTL  
PALSTGLIHONVVVQVLYLIGTSAVVSFAIKMRYVLIILILADARVCCIMMML  
IAQARATLENLVVNAASVAGHGLSLVFPCAAMYTKGRIVPCAAYALGVPFL  
LILALPQAVYAMDREMAASCGAVFVGLILTLSPHYKFLTRPIMVLYQVFTTRFAH  
LQVMIPLNVRGGDAVILITCAIHPELIFTTKILALILGMLVLAQIGITVPEAR  
AHGLIRACMLYKAVAGHYVQWALMKLALITGYDHLTPRDAHAGLNDLAVAE  
PVFSDMETKVTWGPDTAAGDILILGVPASRRRELIHIGPADSLGEGGRILAHIT  
AYSGTRIGLICITSLTGRDNQVEGEVQWSTQSFATCVNGVCTWYTHAGSK  
TLAGRKPTTQMTYNTVDODLVGMQAPGAKSLTPCTGSSDILVITRADVITPVRRG  
DSRGSLSPRPVSYLKSSGGPLCPGSAVGIPLAAVCTRGVAAVDFVPEVSEMTT  
MKSPTVFNSSPPAVPQTFQVAHHAHAPGSGSTVPAVYAAQGYKVLVLPVSVALT  
SFGAYMSKAHGIDPNIKRGVTRTTGAPITVSTYKFLADGGCGSGAVDIIICDECHS

CDs  
  
1..9033  
/product="polyprotein"  
/protein\_id="CAB53095.1"  
/db\_xref="GI:5748511"  
/db\_xref="GOA:Q9Q066"  
/translation="MSTNPKPQRTKRTNRR.POVNKPFGGGQIVGVYL.PRRPRL  
GYRATKTSERSGRGRPIPKAQPRGRMAQGYWPLYNGEGIMAGMULSPRG  
SRPSGPTDPRKSRNLGKVIDLTTCGFDLMGYLPLGAPLGGARLHAGVRLD  
GNVYATGNLPGCSFSIPLIALISCLTIPASAVEVNVGAVHVNDCSNASIVYEAAD  
MIMTPGCGVPCVRENNSRCWALPTLAAANASPTTIRRHVDLILGAALCSAM  
VGDLCGVFLVAQLFTFSPRHEVTDNCISYPGHATGHRAMDMNMNPTALIVY  
SOLIRIPOAVUMVAAHMGVLAGLAVYMGVMAKVLIVMLPAGVDGGRVYCGTM  
AKNTIGITSLSPSGOKIQLVNTNGSHINRTALNCDSLNTGLALPVIHKNSS  
GCEPBMASCSPIDAQCGMGPITTVNESHSDRPYCMHYPAPCGIVPAADCGGVYC  
FTPSVVVGTDRFGVPTPYSGWENETDVLNNTSPGDMNGTANSTGTCKCG  
PPCNIGIGISKTLCTPCDFRGPRPATYTKCGSGPMLIPCLVHYPRIMHPCTVNF  
TIFKRVTVGGVHEHLEAAQWTRGERCNLEDRDSEILSLISTTEMOVLPCSPFTL  
PALSTGLIHONVVVQVLYLIGTSAVVSFAIKMRYVLIILILADARVCCIMMML  
IAQARATLENLVVNAASVAGHGLSLVFPCAAMYTKGRIVPCAAYALGVPFL  
LILALPQAVYAMDREMAASCGAVFVGLILTLSPHYKFLTRPIMVLYQVFTTRFAH  
LQVMIPLNVRGGDAVILITCAIHPELIFTTKILALILGMLVLAQIGITVPEAR  
AHGLIRACMLYKAVAGHYVQWALMKLALITGYDHLTPRDAHAGLNDLAVAE  
PVFSDMETKVTWGPDTAAGDILILGVPASRRRELIHIGPADSLGEGGRILAHIT  
AYSGTRIGLICITSLTGRDNQVEGEVQWSTQSFATCVNGVCTWYTHAGSK  
TLAGRKPTTQMTYNTVDODLVGMQAPGAKSLTPCTGSSDILVITRADVITPVRRG  
DSRGSLSPRPVSYLKSSGGPLCPGSAVGIPLAAVCTRGVAAVDFVPEVSEMTT  
MKSPTVFNSSPPAVPQTFQVAHHAHAPGSGSTVPAVYAAQGYKVLVLPVSVALT  
SFGAYMSKAHGIDPNIKRGVTRTTGAPITVSTYKFLADGGCGSGAVDIIICDECHS

TDSTTILIGITVLDOAETAGARLVVLATATPFGSVTVPHPIIEEVALSSTGEVPEYK  
AIPIETIKGGRHLIFCHSKKKCDELLAALSLGAINAAVAYRGLVSVITGDIYVVA  
TDALMTGFTPDVSDVNDCTVQTVQVPSLPTPTPIETITVPODVAVSQRGRGRG  
RIGIYRPTVTPGERSGMDSVLCEDVACGAWELTPAETSVRLATYTRGILPYOC  
DHLEFWSVFTGLTHIDAHFLSOTKQAGDNFPYVAVOATTCABAQAPSPWDQMKC  
LIRKPLTHGPTPLIRLGAONVQVTTTHPIIKYIMCMSADLSEVITWLVGVLA  
ALAAVCLTTSQSVIVGRVILSGKPAIIPDRREVLYREDEMEBCASHLPYIQGMQLAR  
OPKQALIGLQATKQAAAPVAVESKWTLEAFMAKMMNFISGIQYLAGLSTLPGN  
PAISLWFTASITSPITTOHTLLENILGAVVAAQOLAPPSAASAFVGGAGVAGSI  
GLGKVIDIILAGVAGVAGLVAKMSGEMSTEDIVNLLPILSPALVAVGCA  
ILRRHVGBEAVQWNRLLIPASRGHVSPTHVYPSDAAARVTOILSLITTOCLK  
RLHQMINDCSTPSCGSLRDVMDMTCVLTDFKTVLQSKLPLPGVPFSCQKPK  
GWRGDIMQTTCCGAQITGVHNGSRIRIGPPTCSNTWGTGPPINAYTGPCTPSP  
APNYSRALMRVAAEEVYEVTVGDPEHVTGTTDNVACPCQVPALEFTEVDYALHR  
YAPKCPILREEVFLVNGVYLVGSGLPCEPEPDVAVLTMLDPSHITATPKRL  
KARGPPLIASSASASOLASPLKATCTRHSDPADLIEANILMROCMGNTITRESN  
KVILDSFEPLRADEBEREVSVAELIRSRKPRAMPINARPDTPNPLESWDPDY  
VPVPHGCPPLPPAKAPPIPPRRRRTVLSSTVSAALATATTFGSESSADSGT  
ATAPDPDPSDDAGSDVGSYSNMPLGEGEDPLDPSGWSVSEEGEDVCCMS  
YTMGALITPCAAETKLPINALNSLIRHNLVYATTSRASLRKKKVTDFRLOVLD  
DHRVDVLKEMKAKASTYKAKLISVEBCKLTPPHSARSKPGYAKDVKNLSKAVSHI  
RSWKDULEDETETPIDTTINAKNEVFCQPEKGRKARLIVPDLGVRCEKALYD  
VSTLPLQAVMGSSTGFQISPGQVPELVNTWKSCKCPWGFAYDRCEPSTVTEMDIRV  
TESIYOCDDLPEAROAIRSLTERLYIGPLTNSGQCGYRRASGVLTSSGNTL  
TCYKAAACAAKALQDCTMLVCGDDLVLCESAGTOBDESLAFTEAMTRYAPG  
DPPKREYDLEILITCSSNVSAHADSGKRVYLLTRDPTPLARAMEARTHPVNSWL  
GNITMYAPLTMAEMILMTHPRLSILLAGOKEKALDCOYGAQIPEPLDIPOLIORL  
GLSAFSLHSYSGEINRYASCTRKLGCPPLRVWHRARASVYARLLSQGGRATYCKIL  
FNMAVRTKLKLTPIPAASQLDLSMFVAGVSGDYRSLRSBARPMWMLLILSVG  
GIYLLPNR"  
1..573  
/gene="core"  
/product="core protein"  
574..1149  
/gene="E1"  
/product="glycoprotein E1"  
1150..2238  
/gene="E2"  
/product="glycoprotein E2"  
2239..2427  
/gene="p7"  
/product="p7 peptide"  
2428..3078  
/gene="NS2"  
/product="NS2 protease"  
3079..4971  
/gene="NS3"  
/product="NS3 protease/helicase"  
5134..5916  
/gene="NS4B"  
/product="NS4B protein"  
1..573  
/gene="core"  
574..1149  
/gene="E1"  
1150..2238  
/gene="E2"  
2239..2427  
/gene="p7"  
/product="p7 peptide"  
2428..3078  
/gene="NS2"  
/product="NS2 protease/helicase"  
3079..4971  
/gene="NS3"  
/product="NS3 protease cofactor"  
5134..5916  
/gene="NS4B"  
/product="NS4B protein"  
5917..7256  
/gene="NS5A"  
/product="NS5A protein"  
5917..7256  
/gene="NS5A protein"

```

ORIGIN
Query Match      86.8%; Score 1570.6; DB 14; Length 9033;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

gene      /gene="NS5A"
          /product="NS5A phosphoprotein"
          /size="9030"
          /gene="NS5B"
          /size="9030"
          /gene="NS5B"
          /product="NS5B RNA dependant RNA polymerase"

mat_peptide

1  ATGAGCAGCAATCTCTTAACCTCAAGAAAAACCAACGTAACCAACCGCCGCCACAG 60
Db 1  ATGAGCAGCAATCTCTTAACCTCAAGAAAAACCAACGTAACCAACCGCCGCCACAG 60
QY 61  GACGTCAAGTCCCGGCGGAGTGTAGATCGTTGTGTGAGATTACCTGTTGCGCGCAG 120
Db 61  GACGTCAAGTCCCGGCGGAGTGTAGATCGTTGTGTGAGATTACCTGTTGCGCGCAG 120
QY 121  GGCCTCAGGTTGGGTGCGCGGAGTGAAGACTTCCGAGCGTGCACACTCGTGA 180
Db 121  GGCCTCAGGTTGGGTGCGCGGAGTGAAGACTTCCGAGCGTGCACACTCGTGA 180
QY 181  AGGCGCAACCTATCCCAAGGCTCGCAGCCCGAGGAGGAGGCTGGGCTCAGCCCGG 240
Db 181  AGGCGCAACCTATCCCAAGGCTCGCAGCCCGAGGAGGAGGCTGGGCTCAGCCCGG 240
QY 241  TACCTTGGAGCCCTCTATGAGCAACGAGGAGTGGGAGGAGGAGTGGCTCTGTCACC 300
Db 241  TACCTTGGAGCCCTCTATGAGCAACGAGGAGTGGGAGGAGGAGTGGCTCTGTCACC 300
QY 301  CGTGGCTCCCGGCTAGTGGGAGCCCACTGACCCCGCGGAGTGCCTAATTTGGGT 360
Db 301  CGTGGCTCCCGGCTAGTGGGAGCCCACTGACCCCGCGGAGTGCCTAATTTGGGT 360
QY 361  AAAAGTATGATATCCCTCAATGCGGCTTGCAGACTCAATGGGGTATATTCGGCTGTC 420
Db 361  AAAAGTATGATATCCCTCAATGCGGCTTGCAGACTCAATGGGGTATATTCGGCTGTC 420
QY 421  GGGCGTCCCTTAAGGAGGAGCGTGCAGAGGAGCCCTGGAGCATGCGCTCGAGGAGAC 480
Db 421  GGGCGTCCCTTAAGGAGGAGCGTGCAGAGGAGCCCTGGAGCATGCGCTCGAGGAGAC 480
QY 481  GGGGTGAATATGCAACAGGAAATTTAACCGGTTGCTTTCTATCTTCTTTGGCT 540
Db 481  GGGGTGAATATGCAACAGGAAATTTAACCGGTTGCTTTCTATCTTCTTTGGCT 540
QY 541  TTGCTGTCTGTGTTGACATTCAGCTTCCGCTTAATGAGTGGCAACGTTGCGGAGTC 600
Db 541  TTGCTGTCTGTGTTGACATTCAGCTTCCGCTTAATGAGTGGCAACGTTGCGGAGTC 600
QY 601  TACCATGTACAGCAAGATTGCTCAACTAAGCATGTGTAGAGAGAGGAGCATGATC 660
Db 601  TACCATGTACAGCAAGATTGCTCAACTAAGCATGTGTAGAGAGAGGAGCATGATC 660
QY 661  ATGACACACCCCGGAGTGTGCTCTGTGTCGGGAGAGTAATTCCTCCGCTGCTGGGTA 720
Db 661  ATGACACACCCCGGAGTGTGCTCTGTGTCGGGAGAGTAATTCCTCCGCTGCTGGGTA 720
QY 721  GGGCTCACTTCCACGCTCGCGGCAAGAGAGCAGACATCCCACTGAGACATACGAGC 780
Db 721  GGGCTCACTTCCACGCTCGCGGCAAGAGAGCAGACATCCCACTGAGACATACGAGC 780
QY 781  CAGGTGATTTGCTGTTGGGAGGAGCTTCTGCTCGCTGATGATGAGTGGGAGATCTC 840
Db 781  CAGGTGATTTGCTGTTGGGAGGAGCTTCTGCTCGCTGATGATGAGTGGGAGATCTC 840
QY 841  TGGGATCTGTTTCTCTGCTCTCTAGCTGTCACTTCTGCTCGCGCAGATCAGAGC 900
Db 841  TGGGATCTGTTTCTCTGCTCTCTAGCTGTCACTTCTGCTCGCGCAGATCAGAGC 900

```

```

QY 901  GTACAGAGCTGCAATTTGTTCAATCTATCCCGGACAGTATCAGGTACCCGATGCTTGG 960
Db 901  GTACAGAGCTGCAATTTGTTCAATCTATCCCGGACAGTATCAGGTACCCGATGCTTGG 960
QY 961  GATATGATGATGAATGCTGCTACCTACAGAGCCCTAGTGTATTCGAGCTACTCCGATC 1020
Db 961  GATATGATGATGAATGCTGCTACCTACAGAGCCCTAGTGTATTCGAGCTACTCCGATC 1020
QY 1021  CCACAGCTGTGTGTGACATGTGTGGGAGGAGCCCACTGGGAGTCTGTGGCGGCTTGGC 1080
Db 1021  CCACAGCTGTGTGTGACATGTGTGGGAGGAGCCCACTGGGAGTCTGTGGCGGCTTGGC 1080
QY 1081  TACTATTCATGATGTGGGAACTGGGCTAAGTCTTGTGTATGATGCTATCTTTGGCGGC 1140
Db 1081  TACTATTCATGATGTGGGAACTGGGCTAAGTCTTGTGTATGATGCTATCTTTGGCGGC 1140
QY 1141  GTTGAAGGAGAACTTTACAGCAGAGGAGGAGACACGAGCGGCGCCGACGAGCTTACA 1200
Db 1141  GTTGAAGGAGAACTTTACAGCAGAGGAGGAGACACGAGCGGCGCCGACGAGCTTACA 1200
QY 1201  TCCCTCTTACACCTTGGGCGGCTCAGAAATTCAGCTTGTATTAACCAACGAGAGCTGG 1260
Db 1201  TCCCTCTTACACCTTGGGCGGCTCAGAAATTCAGCTTGTATTAACCAACGAGAGCTGG 1260
QY 1261  CACATCAAGAACTGCTTGAACAGTGAAGTCACTCCCTCAACTGGGTTCTTGGCGCG 1320
Db 1261  CACATCAAGAACTGCTTGAACAGTGAAGTCACTCCCTCAACTGGGTTCTTGGCGCG 1320
QY 1321  CTGTTCTACACGACAGTTCATATGCTGCGGAGTGTCAAGGAGGAGGAGGAGGAGGAG 1380
Db 1321  CTGTTCTACACGACAGTTCATATGCTGCGGAGTGTCAAGGAGGAGGAGGAGGAGGAG 1380
QY 1381  CCCATTGACCAATTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1381  CCCATTGACCAATTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441  GACCAAGAGCCCTATTTGCTGAGCACTACGACCTCAACCGTGTATTCGTGCCGAGT 1500
Db 1441  GACCAAGAGCCCTATTTGCTGAGCACTACGACCTCAACCGTGTATTCGTGCCGAGT 1500
QY 1501  CAGGTGTGTGCTCAGTGTACTGTTCACTCAAGCCCTGTTGTGGGAGGAGGAGGAGT 1560
Db 1501  CAGGTGTGTGCTCAGTGTACTGTTCACTCAAGCCCTGTTGTGGGAGGAGGAGGAGT 1560
QY 1561  CGTTTGGGCGCCCTTACGTAAGATGGGAGTGAATGAACAGGACCTGTGCTTCTTAAC 1620
Db 1561  CGTTTGGGCGCCCTTACGTAAGATGGGAGTGAATGAACAGGACCTGTGCTTCTTAAC 1620
QY 1621  AACAGCGGCGGCAAGGAGCACTGTGCTGCTGTAATGATGATGATGATGATGATGATGAT 1680
Db 1621  AACAGCGGCGGCAAGGAGCACTGTGCTGCTGTAATGATGATGATGATGATGATGATGAT 1680
QY 1681  ACCAAGACGTGTGGGAGCCCGGCTGCAACATCGGAGGAGTGGGCAACACTTTGATC 1740
Db 1681  ACCAAGACGTGTGGGAGCCCGGCTGCAACATCGGAGGAGTGGGCAACACTTTGATC 1740
QY 1741  TGGCCAGAGGAGTCTTCCGGAAGCATCCCGAGGAGCACTTACACCAAAATGCGGTTGGGG 1800
Db 1741  TGGCCAGAGGAGTCTTCCGGAAGCATCCCGAGGAGCACTTACACCAAAATGCGGTTGGGG 1800
QY 1801  CCTTGGTTG 1809
Db 1801  CCTTGGTTG 1809

```

Search completed: February 19, 2005, 07:08:48  
 Job time : 8047.09 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2005, 20:55:25 ; Search time 1012.71 Seconds  
(without alignments)  
10574.404 Million cell updates/sec

Title: US-09-664-363-21\_COPY\_308\_2116

Perfect score: 1809  
Sequence: 1 ATGAGACACCAATCCTTAACCC.....GGCGTTGGGGCCTTGTTG 1809

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1805.8	99.8	2116	2	AAQ12242
2	1572.2	86.9	9605	6	ABK91431
3	1572.2	86.9	9605	6	ABK91424
4	1572.2	86.9	9605	6	ABK91429
5	1572.2	86.9	9605	6	ABK91432
6	1572.2	86.9	9605	6	ABK91411
7	1572.2	86.9	9605	6	ABK91430
8	1572.2	86.9	9605	6	ABK91428
9	1572.2	86.9	9605	6	ABK91425
10	1572.2	86.9	9605	6	ABK91426
11	1572.2	86.9	9605	6	ABK91433
12	1572.2	86.9	9605	6	AAQ25332
13	1572.2	86.9	9605	6	ABK91427
14	1572.2	86.9	11062	3	AAQ25331
15	1570.6	86.8	11076	3	AAQ89655
16	1565.8	86.6	2187	2	ABA03491
17	1565.8	86.6	2540	2	AAQ43889
18	1564.2	86.5	2540	2	AAQ63753
19	1559.4	86.2	2540	2	AAQ29628
20	1556.2	86.0	9587	13	ADR82189

21	1556.2	86.0	9595	2	AAQ24843
22	1556.2	86.0	9595	4	AAQ23492
23	1556.2	86.0	9595	4	AAQ69939
24	1556.2	86.0	9595	12	ADQ36222
25	1556.2	86.0	9595	12	ADQ79396
26	1556.2	86.0	9595	2	AAQ24833
27	1554.6	85.9	3461	2	AAQ64068
28	1554.6	85.9	3461	2	AAQ30386
29	1551.2	85.7	2829	8	AAQ55222
30	1551.2	85.7	2829	2	AAQ60673
31	1549.8	85.7	2187	2	ABA03492
32	1543.4	85.3	3360	2	AAQ10367
33	1543.4	85.3	9413	2	AAQ81559
34	1543.4	85.3	9413	2	AAQ29600
35	1543.4	85.3	9413	6	AAQ25517
36	1543.4	85.3	9413	8	AAQ53723
37	1543.4	85.3	9413	8	AAQ49655
38	1543.4	85.3	9413	10	ADQ88596
39	1540.2	85.1	9405	2	AAQ41345
40	1540.2	85.1	9405	2	AAQ40426
41	1540.2	85.1	9413	2	AAQ80498
42	1540.2	85.1	9611	13	ADQ34713
43	1538.6	85.1	2433	2	AAQ12974
44	1538.6	85.1	2433	10	ADQ55557
45	1538.6	85.1	2433	12	ADQ71139

## ALIGNMENTS

RESULT 1	AAQ12242	standard; DNA; 2116 BP.
ID	AAQ12242	standard; DNA; 2116 BP.
AC	AAQ12242;	
XX		
DT	25-MAR-2003 (revised)	
DT	17-SRP-1991 (first entry)	
XX		
DE	Encodes PT-NANBH viral structural and non-structural proteins.	
XX		
KM	post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.	
XX		
OS	Non-A.	
OS	non-B hepatitis virus.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	308..2116
FT		/*tag= a
XX		
XX	GB239245-A.	
XX		
PD	26-JUN-1991.	
XX		
PF	17-DEC-1990;	90GB-00027250.
XX		
PR	18-DEC-1989;	89GB-00028562.
PR	27-FEB-1990;	90GB-00004414.
PR	03-MAR-1990;	90GB-00004814.
XX		
PA	(WEL) WELLCOME FOUND LTD.	
PA	(HIGH) HIGHFIELD P E.	
XX		
PI	Highfield PE, Rodgers BC, Tedder RS, Barbara JAJ;	
XX		
DR	WPI; 1991-187584/26.	
DR	P-PSDB; AAR12600.	
XX		
PT	Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA	
PT	and antibodies used in diagnostic assays and in vaccines.	
XX		
PS	Claim 10; Page 83-87; 108pp; English.	
XX		

CC This sequence is thought to encode viral structural and non-structural  
CC proteins of the PT-NANBH viral genome which are antigenic. It was  
CC isolated from human serum infections for the virus. See also A042236-41.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX

Sequence 2116 BP; 392 A; 650 C; 624 G; 450 T; 0 U; 0 Other;

Query Match 99.8%; Score 1805.8; DB 2; Length 2116;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1807; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAGCAGCAATCTTAACTTCAAGAAACCAACGTAACCAACCGCCGCCACAG 60
DB 308 ATGAGCAGCAATCTTAACTTCAAGAAACCAACGTAACCAACCGCCGCCACAG 367
QY 61 GACGTCAAGTCCCGGCGGCGGTGTCAGATGCTTGGTGGATTACTCTTGGCCGCGAG 120
DB 368 GACGTCAAGTCCCGGCGGCGGTGTCAGATGCTTGGTGGATTACTCTTGGCCGCGAG 427
QY 121 GGCCTCAGGTTGGGTGTGCGCGCACTAGAGAAAGTCTCCAGCGGTGCAACTCTGTGA 180
DB 428 GGCCTCAGGTTGGGTGTGCGCGCACTAGAGAAAGTCTCCAGCGGTGCAACTCTGTGA 487
QY 181 AGGCGCAACCTATCTCCAGGCTGCGACGCTGAGGCGAGGCGCTGCGCTCAGCCGCG 240
DB 488 AGGCGCAACCTATCTCCAGGCTGCGACGCTGAGGCGAGGCGCTGCGCTCAGCCGCG 547
QY 241 TACCCCTGGGCCCCCTCTATGAGCAAGAGGCGATGAGGCTGGGCGAGAGTGGCTCTGTCACCC 300
DB 548 TACCCCTGGGCCCCCTCTATGAGCAAGAGGCGATGAGGCTGGGCGAGAGTGGCTCTGTCACCC 607
QY 301 CGTGGCTCCCGGCGCTAGTGGGCGCCCACTGACCCCGCGCGTATGCGTGAATTTGGGT 360
DB 608 CGTGGCTCCCGGCGCTAGTGGGCGCCCACTGACCCCGCGCGTATGCGTGAATTTGGGT 667
QY 361 AAAGTCATGATACCTCTCATAGCGGCTTGCACGACTCATGAGGTAATTCGCTCTGCT 420
DB 668 AAAGTCATGATACCTCTCATAGCGGCTTGCACGACTCATGAGGTAATTCGCTCTGCT 727
QY 421 GCGCGTCCCTTATGAGGCGCGCTGCAAGGCGCTGGGCGATGAGGCTCCGGGTTTGTGAAGAG 480
DB 728 GCGCGTCCCTTATGAGGCGCGCTGCAAGGCGCTGGGCGATGAGGCTCCGGGTTTGTGAAGAG 787
QY 481 GCGCGTAACTATGACAGAGGAATTTACCGGTTGCTTCTCTATCTTCTCTTGGCT 540
DB 788 GCGCGTAACTATGACAGAGGAATTTACCGGTTGCTTCTCTATCTTCTCTTGGCT 847
QY 541 TTGCTGCTCTGTTTGAACATTCAGCTTCGCTTATGAAGTGCAGCAAGTGTCCGGGATC 600
DB 848 TTGCTGCTCTGTTTGAACATTCAGCTTCGCTTATGAAGTGCAGCAAGTGTCCGGGATC 907
QY 601 TACCAATGACGAAGATGCTGTCACATCAAGCATGCTGTAGAGACAGCGGACATGATC 660
DB 908 TACCAATGACGAAGATGCTGTCACATCAAGCATGCTGTAGAGACAGCGGACATGATC 967
QY 661 ATGCAACACCCCGGCGGTGTGCGCTGTGTCCGAGAGGTAATTCCTCCGCTGTGGGTA 720
DB 968 ATGCAACACCCCGGCGGTGTGCGCTGTGTCCGAGAGGTAATTCCTCCGCTGTGGGTA 1027
QY 721 GCGCTCACTCCACGCTGCGGCGCAAGAGCGCAAGATCCCACTGCGCAATACGACGC 780
DB 1028 GCGCTCACTCCACGCTGCGGCGCAAGAGCGCAAGATCCCACTGCGCAATACGACGC 1087
QY 781 CACGTGATGTTGCTGTTGGGCGCGGCTTCTGCTCGCATATGATAGTGGGGATCTC 840
DB 1088 CACGTGATGTTGCTGTTGGGCGCGGCTTCTGCTCGCATATGATAGTGGGGATCTC 1147
QY 841 TCGGATCTGTTTCTCTCTCTCTCAGCTGTTCACCTTCTCGCTCGCGCATCAGAGC 900
DB 1148 TCGGATCTGTTTCTCTCTCTCTCAGCTGTTCACCTTCTCGCTCGCGCATCAGAGC 1207
QY 901 GTACAGACTGCAATTTGCAATCTATCCCGGCGACGATACAGTCAACGCGATGCTTGG 960
DB 1148 TCGGATCTGTTTCTCTCTCTCTCAGCTGTTCACCTTCTCGCTCGCGCATCAGAGC 1207
```

```
DB 1208 GTACAGACTGCAATTTGCAATCTATCCCGGCGACGATACAGTCAACCGCATGCTTGG 1267
QY 961 GATATGATGATGAACTGTGTCACTTACAGCAGCCCTAGTGTATCGACGTACTCCGATC 1020
DB 1268 GATATGATGATGAACTGTGTCACTTACAGCAGCCCTAGTGTATCGACGTACTCCGATC 1327
QY 1021 CCACAGCTGTCTGTGACATGATGTGCGGGGCCCACTGGGAGTCTTGGCGGCTTGGCC 1080
DB 1328 CCACAGCTGTCTGTGACATGATGTGCGGGGCCCACTGGGAGTCTTGGCGGCTTGGCC 1387
QY 1081 TACTATTCAATGATGTGGGAACTGGGCTTAAGGCTTGTGTGTGATGCTACTCTTGGCCGCG 1140
DB 1388 TACTATTCAATGATGTGGGAACTGGGCTTAAGGCTTGTGTGTGATGCTACTCTTGGCCGCG 1447
QY 1141 GTTGAACGGGGAACCTTACAGCAAGGGGGGACACACGCGCGCGGCCACGAGGCTTACA 1200
DB 1448 GTTGAACGGGGAACCTTACAGCAAGGGGGGACACACGCGCGCGGCCACGAGGCTTACA 1507
QY 1201 TCCCTCTTCAACCTGTGGCGCGCTCAGAAATCCAGCTTGTAAACACCAACGAGGAGCTGG 1260
DB 1508 TCCCTCTTCAACCTGTGGCGCGCTCAGAAATCCAGCTTGTAAACACCAACGAGGAGCTGG 1567
QY 1261 CACATCAACAGAACTGCTTGAATGCAATGCTCTCTCCAACTGGGTTCTTGGCCGCG 1320
DB 1568 CACATCAACAGAACTGCTTGAATGCAATGCTCTCTCCAACTGGGTTCTTGGCCGCG 1627
QY 1321 CTGTTCTTACAGCAAGGTTCAATGCGTCCGAGATGCTCAGAGCGCATGAGCGCTGCGC 1380
DB 1628 CTGTTCTTACAGCAAGGTTCAATGCGTCCGAGATGCTCAGAGCGCATGAGCGCTGCGC 1687
QY 1381 CCCATTGACCAATTCGATCAGAGGCTGGGCTCCATCACTTATTAATGAGTCCACGAGCTTG 1440
DB 1688 CCCATTGACCAATTCGATCAGAGGCTGGGCTCCATCACTTATTAATGAGTCCACGAGCTTG 1747
QY 1441 GACCAAGGCGCTTATGCTGCGACTACGACACTCAACCGTGTGTATGTGCTCCGCTTG 1500
DB 1748 GACCAAGGCGCTTATGCTGCGACTACGACACTCAACCGTGTGTATGTGCTCCGCTTG 1807
QY 1501 CAGGTGTGCGCCAGTGTACTGTTTCACTGCAAGCCCTGTTGTGTGGGAGCAGCGAT 1560
DB 1808 CAGGTGTGCGCCAGTGTACTGTTTCACTGCAAGCCCTGTTGTGTGGGAGCAGCGAT 1867
QY 1561 CGTTTGGCGCGCCCTTACGATCAGATGAGGTGATGATGAGACGAGCGTCTTCTCAAC 1620
DB 1868 CGTTTGGCGCGCCCTTACGATCAGATGAGGTGATGATGAGACGAGCGTCTTCTCAAC 1927
QY 1621 AACACGCGCGCGCCACGCGGCGAACTGTTGCGGCTGTACATGATGATGATGACACGCGGTTTC 1680
DB 1928 AACACGCGCGCGCCACGCGGCGAACTGTTGCGGCTGTACATGATGATGATGACACGCGGTTTC 1987
QY 1681 ACCAAGAGTGTGGGGGCGCCCGCGCAACATCGGGGGGCTCGGCAACCACTTTGATC 1740
DB 1988 ACCAAGAGTGTGGGGGCGCCCGCGCAACATCGGGGGGCTCGGCAACCACTTTGATC 2047
QY 1741 TGCCCAACGAGCTGCTTCCGAGACATCCGAGGCACTTACACCAATGCGTTCGGGG 1800
DB 2048 TGCCCAACGAGCTGCTTCCGAGACATCCGAGGCACTTACACCAATGCGTTCGGGG 2107
QY 1801 CTTTGGTTG 1809
DB 2108 CTTTGGTTG 2116
```

## RESULT 2

ABK91431  
ID ABK91431 standard; DNA; 9605 BP.

ABK91431;

15-NOV-2002 (first entry)

Hepatitis C virus Con 1 isolate DNA mutant 8.

KM HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KM internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
XX  
OS Hepatitis C virus.  
OS Synthetic.  
FH  
XX  
FH Key Location/Qualifiers  
CDS 342..9374  
FT /tag= a  
FT /product= "HCV polyprotein"  
FT /note= "The polyprotein consists of the Core, E1, E2, p7,  
FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
FT rebase(6934,T)  
FT /tag= b  
FT mutation  
PN WO200259321-A2.  
XX  
XX 01-AUG-2002;  
PD 16-JAN-2002; 2002WO-EP000526.  
XX  
XX 23-JAN-2001; 2001US-0263479P.  
XX  
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.  
XX De Francesco R, Migliaccio G, Paonessa G;  
XX  
XX WPI; 2002-599793/64.  
XX  
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX  
XX Claim 9; Page; 69pp; English.  
XX  
XX The invention relates to nucleic acid molecules comprising altered HCV  
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
CC internal ribosome entry site (IRES) region coding for one or more NS3,  
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
CC are detailed in the specification. Also included are (1) an expression  
CC vector comprising a nucleotide sequence coding for the altered nucleic  
CC acid, which is transcriptionally coupled to an exogenous promoter; (2) a  
CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
CC; (3) a recombinant cell produced by introducing into a human hepatoma  
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC protein, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
CC mutant of the invention. Note: The present sequence is not shown in the  
CC specification but was created by the indexer using the HCV sequence  
CC appearing as ABK91411 and the information in Claim 9  
XX  
SQ Sequence 9605 BP; 1910 A; 2884 C; 2733 G; 2078 T; 0 U; 0 Other;  
Query Match 86.9%; Score 1572.2; DB 6; Length 9605;  
Best Local Similarity 91.8%; Pred. No. 0;  
Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 121 GGCCCCAGGTTGGGTGCGCGGACTAGAAAGACTTCCGAGCGGTGCGAAGCTCGTGA 180  
DB 462 GGCCCCAGGTTGGGTGCGCGGACTAGAAAGACTTCCGAGCGGTGCGAAGCTCGTGA 521  
QY 181 AGCGCAACACTATCCCAAGGCTGCGCAGCCGAGGCGAGGCTCGAGCCCGAG 240  
DB 522 AGCGCAACACTATCCCAAGGCTGCGCAGCCGAGGCGAGGCTCGAGCCCGAG 581  
QY 241 TACCTTGGCCCTCTATGAGCAACGAGGCGATGAGGCTGAGAGTGGCTCTGTACCC 300  
DB 582 TACCTTGGCCCTCTATGAGCAACGAGGCGATGAGGCTGAGAGTGGCTCTGTACCC 641  
QY 301 CGTGGCTCCCGGCTTATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
DB 642 CGTGGCTCCCGGCTTATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 701  
QY 361 AAGTCATGATACCTCTCATGCGGCTTCCGCGAAGCTGAGGCTATCTCGCTCGTC 420  
DB 702 AAGTCATGATACCTCTCATGCGGCTTCCGCGAAGCTGAGGCTATCTCGCTCGTC 761  
QY 421 GGGCTCCCTTAAAGGCGGCTGCGCAGGCGCTGAGCGCATGAGGCTCGGAGTTCGAGGAC 480  
DB 762 GGGCTCCCTTAAAGGCGGCTGCGCAGGCGCTGAGCGCATGAGGCTCGGAGTTCGAGGAC 821  
QY 481 GGGCTCCCTTAAAGGCGGCTGCGCAGGCGCTGAGCGCATGAGGCTCGGAGTTCGAGGAC 540  
DB 822 GGGCTCCCTTAAAGGCGGCTGCGCAGGCGCTGAGCGCATGAGGCTCGGAGTTCGAGGAC 881  
QY 541 TTGCTGCTGTTTGAACATTCAGCTTCGCTTATGAGTGGCAAGCTGTCGGAGTTC 600  
DB 882 TTGCTGCTGTTTGAACATTCAGCTTCGCTTATGAGTGGCAAGCTGTCGGAGTTC 941  
QY 601 TACATGTACGAGGATGCTTCACTCAAGCTGATGATGAGGACAGGACATGATC 660  
DB 942 TACATGTACGAGGATGCTTCACTCAAGCTGATGATGAGGACAGGACATGATC 1001  
QY 661 ATGCAACCCCGGGGTGTCGCTGTCGCGGAGGATATTCCTCCGCTGCTGGGATC 720  
DB 1002 ATGCAACCCCGGGGTGTCGCTGTCGCGGAGGATATTCCTCCGCTGCTGGGATC 1061  
QY 721 GGGCTCACTCCAGCGCTGCGCAGGCGCAGGCGATCCGATCGAATATGAGGAC 780  
DB 1062 GGGCTCACTCCAGCGCTGCGCAGGCGCAGGCGATCCGATCGAATATGAGGAC 1121  
QY 781 CAGTGTGATTTGCTGTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 1122 CAGTGTGATTTGCTGTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181  
QY 841 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
DB 1182 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241  
QY 901 GTCAGAGATGCAATGTTCAATCTATCCGCGCAGTATCAAGTACCGCATGAGTGG 960  
DB 1242 GTCAGAGATGCAATGTTCAATCTATCCGCGCAGTATCAAGTACCGCATGAGTGG 1301  
QY 961 GATATGATGATGAATGTCACCTTACAGAGGCTTATGAGTATGAGTATGAGTATGAGT 1020  
DB 1302 GATATGATGATGAATGTCACCTTACAGAGGCTTATGAGTATGAGTATGAGTATGAGT 1361  
QY 1021 CCAAGAGCTGCTGAGCAATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080  
DB 1362 CCAAGAGCTGCTGAGCAATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1421  
QY 1081 TACTATTCATGATGAGGCAATGAGGCTTATGAGTATGAGTATGAGTATGAGTATGAGT 1140  
DB 1422 TACTATTCATGATGAGGCAATGAGGCTTATGAGTATGAGTATGAGTATGAGTATGAGT 1481  
QY 1141 GTTGAAGGAGGCACTTATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200  
DB 1482 GTTGAAGGAGGCACTTATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1541

Oy		1201	CCCCCTTACACCCCGGGGCGGGCTCAGAAAAATCCAGCTTTGAATAACAACAAGGACGCCTGG	1286
Dd		1542	TCCCTCTTTTCAACCCGGGTCATGCCAGAAAATCCAGCTTTGTAAACCAACAAGGACGCCTGG	1601
Oy		1261	CACATCAACAAATACTGACTTGAACTGCATGCACTCCCTCAAACAGGGATTCTTTCGCSGC	1320
Dd		1602	CACATCAACAAAGAATGCGCCCTGAACTGCATTAAGTACTCCCTCAAACAGGGATTCTTTCGCSGC	1661
Oy		1321	CTGTTCTTACACGACAGGTTCAATGCGTCCGAGTGTCTCAGAGCGCATGGCAGCTGCCGC	1380
Dd		1662	CTGTTCTTACGTCGACAAAGTTCAAACTCATCTGGATGTCGCCAGAGCGCATGGCAGCTGCCAGC	1721
Oy		1361	CCCAATTGACCAAGTTGATCAGAGGGGGGGGTCGCCAATTAATTAAATGAATGCCACGCGCTTG	1441
Dd		1722	CCCAATGACGCGGTTGCTTCAGAGGGGGGGGCCATCACTTAACAAATGAATGATCACAGCTTCG	1781
Oy		1441	GACCAGAGAGCCCTAATTGCTTGSCACATACGACACTCAACCGTAGTGATGTCGCCGCGCTTG	1500
Dd		1782	GACCAGAGAGCCCTTAATTGTTATGTGGCACTACGCCAACCCCGGCGGTGCTGATACCGCGCGC	1841
Oy		1501	CAGGTGTGTGACCCCAAGTACTGTTTCACTCCAAAGCCCTGTTGTGTGGTGGGACGACCGAT	1560
Dd		1842	CAGGTGTGTGTGATCAGTGTACTGCTTCAACCCCAAGCCCTGTCGTGTGTGGGACGACCGAC	1901
Oy		1561	CGTTTTGCGCGCCCTCCTACGTAAGATGGGGTGAAGATGAGAGGAGCGTCTGCTTCTCAAC	1620
Dd		1902	CGGTTTCGCGCGTCCCTACCTACAGTTGGGGGGGAGATGAGACGGACGTCCTCTCTTTTAAC	1961
Oy		1621	AACACGCGGGCGGCCCAACGGGGGCACTGCTTCGGCTGTATAGATGAATGACACGGCGGTTTC	1680
Dd		1962	AACACGCGGGCGGCCGACAGGCACTGGTTTGCTGTATAGATGAATGATGACCTGGGTTTC	2021
Oy		1681	ACCAAGACGTGTGGGGGGCCCCCGGTGCAAACTCGGGGGGGTTCGGCAACAACACTTTGATC	1740
Dd		2022	ACCAAGACGTGTGGGGGGGGCCCCCGGTGTAACATCGGGGGGATTCGGCAATTAACCTTTGACC	2081
Oy		1741	TGCCCCACGACACTGTTCTCCGGAAGCATCCCGAGGCACTTACACCAATGCGGTTCCGGGG	1800
Dd		2082	TGCCCCACGACACTGCTTCCGGGAAGCACCCCGAGGCCACTTACACCAAGTGTGTTCGGGG	2141
Oy		1801	CGTTGGTTG 1809 	
Dd		2142	CGTTGGTTG 2150 	
<b>RESULT 3</b>				
ID	ABK91424	standard; DNA; 9605 BP.		
XX AC	ABK91424;			
XX DT	15-NOV-2002	(first entry)		
XX DE	Hepatitis C virus Con 1 isolate DNA mutant 1.			
XX KW	HCV, d8; Con 1; adaptive mutation; liver failure; cirrhosis; hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV; internal ribosome entry site; IRES; NS5a; HCV replication; mutant.			
OS	Hepatitis C virus.			
XX OS	Synthetic.			
XX FH	Key	Location/Qualifiers		
FT CDS		342..9374		
FT FT		/*tag= a		
FT FT		/product= "HCV polypeptide"		
FT FT		/note= "The polypeptide consists of the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"		
FT FT	mutation	replace(3625,G) /*tag= b		
XX PN	WO200259321-A2.			

PD		01-AUG-2002.	
XX			
PF		16-JAN-2002; 2002WO-EP000526.	
XX			
PR		23-JAN-2001; 2001US-0263479P.	
XX			
PA		(RICE-) 1ST RICERCHE BIOL MOLECOLARE ANEBLETITI.	
XX			
P1		De Francesco R, Migliaccio G, Paonessa G;	
DR		WPI; 2002-599793/64.	
XX			
PT		New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV	
PT		NS5 encoding region, or encephalomyocarditis virus (EMCV) internal	
PT		ribosome entry site (IRES) region, useful in studying HCV replication and	
PT		expression.	
PS			
XX		Claim 9; Page: 69pp; English.	
CC			
CC		The invention relates to nucleic acid molecules comprising altered HCV	
CC		NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)	
CC		internal ribosome entry site (IRES) region coding for one or more NS3,	
CC		NS5A, or EMCV IRES mutations, respectively. The location of the mutations	
CC		are detailed in the specification. Also included are (1) an expression	
CC		vector comprising a nucleotide sequence coding for the altered nucleic	
CC		acids, which is transcriptionally coupled to an exogenous promoter; (2) a	
CC		recombinant cell human hepatoma cell comprising the altered nucleic acids	
CC		; (3) a recombinant cell produced by introducing into a human hepatoma	
CC		cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)	
CC		replicon enhanced cell or which containing a functional HCV replicon; (5)	
CC		an HCV replicon enhanced cells made in the method, and (6) measuring the	
CC		ability of a compound to affect HCV activity. The HCV replicons and HCV	
CC		replicon enhanced cells are useful in studying HCV replication and	
CC		expression, and HCV and host cell interactions, producing HCV RNA and	
CC		proteins, and providing a system for measuring the ability of a compound	
CC		to modulate one or more HCV activities e.g. to discover drugs which may	
CC		treat HCV mediated diseases such as liver failure, cirrhosis and	
CC		hepatocellular carcinoma. The present sequence is an HCV replicon Con 1	
CC		mutant of the invention. Note: The present sequence is not shown in the	
CC		specification but was created by the indexer using the HCV sequence	
CC		appearing as ABK01411 and the information in Claim 9	
SQ			
	Sequence 9605 BP, 1910 A; 2884 C; 2732 G; 2079 T; 0 U; 0 Other;		
	Query Match	86.9%; Score 1572.2; DB 6; Length 9605;	
	Best Local Similarity	91.8%; Pred. No. 0;	
	Matches 1661; Conservative	0; Mismatches 148; Indels	0; Gaps
QY		1 ATGAGCAGCAATCTCTTAACCTTCAAGAAGAAAAACAACCTAACCATCACCGCGGCCACAG	60
DB		342 ATGAGCAGCAAGAACTCTTAACCTTCAAGAAGAAAAACAACCTAACCATCACCGCGGCCACAG	401
QY		61 GACGTCAAGTTTCGGGGGGGGTGATGCATGCTGTGTGAAGTTTACTGTGGCGGCGAGG	120
DB		402 GACGTCAAGTTTCGGGGGGGGTGATGCATGCTGTGTGAAGTTTACTGTGGCGGCGAGG	461
QY		121 GGCCCCAGAGTTGGGTGTGTGCGCGCACTGAGAAAGCTTTCGACGGTTCGCAACTTGTGGA	180
DB		462 GGCCCCAGAGTTGGGTGTGTGCGCGCACTGAGAAAGCTTTCGACGGTTCGCAACTTGTGGA	521
QY		181 AGGCGACAACTTATCCCAAGGCTTGCCAGGCCCGAGGGCAGGGCTCGAGCCGGG	240
DB		522 AGGCGACAACTTATCCCAAGGCTTGCCAGGCCCGAGGGCAGGGCTCGAGCCGGG	581
QY		241 TACCCTTGGCCCTTATGGAAGAAGAGGCAATGGGGTGGGAGATATGGCTCTGTGCACCC	300
DB		582 TACCCTTGGCCCTTATGGAAGAAGAGGCTTGGGGTGGGAGATATGGCTCTGTGCACCC	641
QY		301 CGTAGCTCCCGAGCTAGTTGGGGGCCCACTACCCCAGGCGTAGGTTCGTAATTGGGT	360
DB		642 CGTAGCTCCCGAGCTAGTTGGGGGCCCACTACCCCAGGCGTAGGTTCGTAATTGGGT	701
QY		361 AAAGCATCATGACCTTCATATGGGGCTTGGCGCACTTCATGGGGGTACATTCGCTGCT	420

[illegible]

Db	1782	GACCAAGAGGCTTATTGTGGCACTACGCAACCCGGACGTGCGGTATCGTACCCGGCGG	1841
Qy	1501	CAGGTGTGTGGCCCAAGTGTACTGTTTCACTTCGAAGCCCTGTGTGTGGGAGACGAT	1560
Db	1842	CAGGTGTGTGTCCAGTGTACTGCTTCAACCCCAAGCCCTGTGTGTGGGAGACGAC	1901
Qy	1561	CGTTTCGGCGGCCCTACGTACAGATGGGGGTGAGAAATGAGACGAGCTGTCTTCAAC	1620
Db	1902	CGTTTCGGCGGCCCTACGTACAGATGGGGGTGAGAAATGAGACGAGCTGTCTTCAAC	1961
Qy	1621	AACACGGGCGCCGCAACGAGGCACTGGTTCGGCTGTACATGATGAATAGCACCGGTTT	1680
Db	1962	AACACGGGCGCCGCAACGAGGCACTGGTTCGGCTGTACATGATGAATAGCACCGGTTT	2021
Qy	1681	ACCAAGACGTGTGGGGGCCCCCGTGCACATCGGGGGGTGGCGAACAACCTTTGATC	1740
Db	2022	ACCAAGACGTGTGGGGGCCCCCGTGCACATCGGGGGGTGGCGAACAACCTTTGATC	2081
Qy	1741	TGCCCCACGCACTGCTTCCGGAAACATCCGAGGCACTTAACCAATCGGTTCCGGG	1800
Db	2082	TGCCCCACGCACTGCTTCCGGAAACATCCGAGGCACTTAACCAAGTGTGTGGGG	2141
Qy	1801	CCTTGGTTG 1809	
Db	2142	CCTTGGTTG 2150	
RESULT 4			
ABK91429	ABK91429 standard; DNA; 9605 BP.		
XX AC	ABK91429;		
XX AC	15-NOV-2002 (first entry)		
XX DE	Hepatitis C virus Con 1 isolate DNA mutant 6.		
XX KW	HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;		
XX KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;		
XX KW	internal ribosome entry site; IRES; NS5A; HCV replication; mutant.		
XX OS	Hepatitis C virus.		
XX OS	Synthetic.		
XX FT	Key	Location/Qualifiers	
XX FT	CDS	342..9374	
XX FT		/*tag= a	
XX FT		/product= "HCV polyprotein"	
XX FT		/note= "The polyprotein consists of the Core, E1, E2, P7,	
XX FT	mutation	NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"	
XX FT		replace(6859,C)	
XX FT		/*tag= b	
XX FN	WO200259321-A2.		
XX PD	01-AUG-2002.		
XX PF	16-JAN-2002; 2002MO-EP000526.		
XX PR	23-JAN-2001; 2001US-0263479P.		
XX PA	(RICE-) IST RICERCH BIOL MOLECOLARE ANGELETTI.		
XX PI	De Francesco R, Migliaiaccio G, Paonessa G;		
XX DR	WPI; 2002-599793/64.		
XX PT	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV		
XX PT	NS5 encoding region, or encephalomyocarditis virus (EMCV) internal		
XX PT	ribosome entry site (IRES) region, useful in studying HCV replication and		
XX PT	expression.		
XX PS	Claim 9; Page: 69pp; English.		

XX The invention relates to nucleic acid molecules comprising altered HCV  
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
CC internal ribosome entry site (IRES) region coding for one or more NS3,  
CC NS5a, or EMCV IRES mutations, respectively. The location of the mutations  
CC are detailed in the specification. Also included are (1) an expression  
CC vector comprising a nucleotide sequence coding for an altered nucleic  
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC proteins, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
CC mutant of the invention. Note: The present sequence is not shown in the  
CC specification but was created by the indexer using the HCV sequence  
CC appearing as ABK91411 and the information in Claim 9

SQ Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 U; 0 Other;

Query Match 86.9%; Score 1572.2; DB 6; Length 9605;

Best Local Similarity 91.8%; Pred. No. 0;

Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1 ATGAGCAGATCTTAAACCTTAAGAAAACCAACGTAACACACCGCCGACAG 60  
DB 342 ATGACACAGAACTCTAAACCTCAAGAAAAACCAACGTAACACACCGCCGACAG 401  
QY 61 GACGTCAGATCCCGGCGGCGGTGAGATCGTTGGAGATTACTGTGGCGGCGAG 120  
DB 402 GACGTCAGATCCCGGCGGCGGTGAGATCGTTGGAGATTACTGTGGCGGCGAG 461  
QY 121 GCGCCAGGTTGGTGTGCGCGCATAGAGAGACTTCCGAGCGGTCCGACCTGTGGA 180  
DB 462 GCGCCAGGTTGGTGTGCGCGCATAGAGAGACTTCCGAGCGGTCCGACCTGTGGA 521  
QY 181 AGCGCAACAACCTATCCCAAGGCTGCGACCGCGAGGCGCTGCGCTCAGCCGAG 240  
DB 522 AGCGCAACAACCTATCCCAAGGCTGCGACCGCGAGGCGCTGCGCTCAGCCGAG 581  
QY 241 TACCCCTGGCGCCTTATAGGCAACGAGGCGATGGGTGGCGAGATGGCTCTGTACCC 300  
DB 582 TACCCCTGGCGCCTTATAGGCAACGAGGCGATGGGTGGCGAGATGGCTCTGTACCC 641  
QY 301 CGTGGCTCCGCGCTAGTTGGGCGCCCACTGACCCCGCGCTGAGGTCAATTGGAGT 360  
DB 642 CGTGGCTCCGCGCTAGTTGGGCGCCCACTGACCCCGCGCTGAGGTCAATTGGAGT 701  
QY 361 AAAGTCATGATACCTTCATGCGGCTTGGCGACCTTCATGAGGATCAATTCGCTGCT 420  
DB 702 AAAGTCATGATACCTTCATGCGGCTTGGCGACCTTCATGAGGATCAATTCGCTGCT 761  
QY 421 GCGGCTCCCTTATGAGGCGCGCTGCGAGGCGCTGGGCGCATGGGTCCGGGTTTGAAGAC 480  
DB 762 GCGGCTCCCTTATGAGGCGCGCTGCGAGGCGCTGGGCGCATGGGTCCGGGTTTGAAGAC 821  
QY 481 GCGGTCATGATGACAGGGAATTACCGGTTCTCTTCTATCTTCTCTTCTGAGCT 540  
DB 822 GCGGTCATGATGACAGGGAATTACCGGTTCTCTTCTATCTTCTCTTCTGAGCT 881  
QY 541 TTGCTGTCCTGTTTGAACATTCAGCTTCGCTTATGAAATGCGGACGCTGTCGGAGTC 600  
DB 882 TTGCTGTCCTGTTTGAACATTCAGCTTCGCTTATGAAATGCGGACGCTGTCGGAGTC 941  
QY 601 TACCATGTCAGAAAGATGTCCTCAACATCAAGCATCTGTAACGACACGCAATGATC 660  
DB 942 TACCATGTCAGAAAGATGTCCTCAACATCAAGCATCTGTAACGACACGCAATGATC 1001

QY 661 ATGACACCCCGGGTGTGTGCTCCGTGTCCGGAGGATTAATCTCCCGTGTGAGTA 720  
DB 1002 ATGATATCCCGGGGTGTGTGCTCCGTGTGGGAGAAACAATCTCTCCGCTGTGGGTA 1061  
QY 721 GCGTCACTCCACGCTCGCGGCGCAAGAGACGCGACATCCCACTGCGCAATTCGACGC 780  
DB 1062 GCGTCACTCCACGCTCGCGGCGCAAGAGACGCGACATCCCACTGCGCAATTCGACGC 1121  
QY 781 CAGCGCATTTGCTGTGTGGGCGGCGCTTCTGCTCCGTATGTAACGTGGGAGATCTC 840  
DB 1122 CATGTCATTTGCTGTGTGGGCGGCGCTCTCTGTCTCGTATGTAACGTGGGAGATCTC 1181  
QY 841 TCGGATCTGTTTCTCTGTCTCTCAGCTGTTCACCTTCTCGCCCGCAATCAGACG 900  
DB 1182 TCGGATCTGTTTCTCTGTCTCTCAGCTGTTCACCTTCTCGCCCGCAAGAGCA 1241  
QY 901 GTACAGACTGCAATTTGTAATCTATCCGCGCAAGTATCAGGTCAACCGATGCTTGG 960  
DB 1242 GTACAGACTGCAATTTGTAATCTATCCGCGCAAGTATCAGGTCAACCGATGCTTGG 1301  
QY 961 GATATGATGAACTGTGCTCACTACAGACGCTTGTGATGAGCACTCTCGGATC 1020  
DB 1302 GATATGATGAACTGTGCTCACTACAGACGCTTGTGATGAGCACTCTCGGATC 1361  
QY 1021 CCACAAGCTGTGTCGATGATGAGCGGCGGCGCACTGCGGAGTCTGCGCGCTTGGC 1080  
DB 1362 CCACAAGCTGTGTCGATGATGAGCGGCGGCGCACTGCGGAGTCTGCGCGCTTGGC 1421  
QY 1081 TACTATTCATGTTGGGGAATCGGCGCTAGGTTCTTGTGTGATGATCTTTTGGCGGC 1140  
DB 1422 TACTATTCATGTTGGGGAATCGGCGCTAGGTTCTTGTGTGATGATCTTTTGGCGGC 1481  
QY 1141 GTTACCGGGGAACCTTACAGACAGGGGGGACACACGCGCGCGGCCACAGGGCTTACA 1200  
DB 1482 GTTACCGGGGAACCTTATGTAACAGGGGGGACATGCGCAAAAACCTTGGGATTTACG 1541  
QY 1201 TCCCTTTTCAACCTTGGCGCGCTCAGAAAATCCAGCTTTTAAACCAACGAGGAGCTG 1260  
DB 1542 TCCCTTTTCAACCTTGGCGCGCTCAGAAAATCCAGCTTTTAAACCAACGAGGAGCTG 1601  
QY 1261 CACATCAACAGAACTGCTTGAATCTGCAATGATCTCTCCAACTGGGTTCTTGGCGCG 1320  
DB 1602 CACATCAACAGAACTGCTTGAATCTGCAATGATCTCTCCAACTGGGTTCTTGGCGCG 1661  
QY 1321 CTGTTTCAACAGCAAGGTTCAATGCGTCCGAGTCTCAGAGCGCATAGGCGAGCGCGC 1380  
DB 1662 CTGTTTCAACAGCAAGGTTCAATGCGTCCGAGTCTCAGAGCGCATAGGCGAGCGC 1721  
QY 1381 CCCATTGACAGTTTCAATGAGGAGTGGGCTCCATCACTTATTAATGATCCACGCGCTTG 1440  
DB 1722 CCCATTGACAGGTTTCAATGAGGAGTGGGCTCCATCACTTATTAATGATCCACGCGCT 1781  
QY 1441 GACCAAGGCGCTTATTTGTGCGACTAGACACCTCAACCGTGTGATCTGCGCGCTTG 1500  
DB 1782 GACCAAGGCGCTTATTTGTGCGACTAGACACCTCAACCGTGTGATCTGCGCGCTTG 1841  
QY 1501 CAGGTGTGCGCCAGTGTACTGTTTCACTCAAGCCCTGTTGTGGGGAGACAGCAT 1560  
DB 1842 CAGGTGTGCGCCAGTGTACTGTTTCACTCAAGCCCTGTTGTGGGGAGACAGCAT 1901  
QY 1561 CGTTTGGCGCGCTTCACTAGATGAGTGGGAGTGAATGAGACGAGCTGTGCTTCTCAAC 1620  
DB 1902 CGTTTGGCGCGCTTCACTAGATGAGTGGGAGTGAATGAGACGAGCTGTGCTTCTCAAC 1961  
QY 1621 AACACGGGCGCGCAAGGCGCACTGTTGCGCTGTACATGATGAATAGCACGGGCTTC 1680  
DB 1962 AACACGGGCGCGCAAGGCGCACTGTTGCGCTGTACATGATGAATAGCACGGGCTTC 2021  
QY 1681 AACCAAGCGTGTGGGGCGCGCGCGTGAACAATCGGGGGGGTGGCAACAACATTTGATC 1740  
DB 2022 AACCAAGCGTGTGGGGCGCGCGCGTGAACAATCGGGGGGGTGGCAACAACATTTGATC 2081



QY 1741 TGCCGACGAGACTGCTTCCGGAAGCATCCGAGGCGCACTTACCAAAATGCGGTGGGG 1800  
 DB 2082 TGCCCCAGACACTGCTTCCGGAAGCATCCGAGGCGCACTTACCAAAATGCTGGGG 2141  
 QY 1801 CCTTGCTTG 1809  
 DB 2142 CCTTGCTTG 2150

RESULT 5  
 ABR91432  
 ID ABR91432 standard; DNA; 9605 BP.  
 AC ABR91432;  
 DT 15-NOV-2002 (first entry)  
 DE Hepatitis C virus Con 1 isolate DNA mutant 9.  
 KM HCV; de; Con 1; adaptive mutation; liver failure; cirrhosis;  
 KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KM internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
 OS Hepatitis C virus.  
 OS Synthetic.

Key Location/Qualifiers  
 CDS 342..9374  
 FT /\*tag= a  
 FT /product= "HCV polyprotein"  
 FT /note= "The polyprotein consists of the Core, E1, E2, p7,  
 FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
 FT mutation replace(6936,G)  
 FT /\*tag= b

MO200259321.A2.  
 PD 01-AUG-2002;  
 PF 16-JAN-2002; 2002WO-BE000526.  
 PR 23-JAN-2001; 2001US-0263479P.  
 PA (RICE-) IST-RICERCHIE BLO MOLECOLARE ANGELETTI.  
 PI De Francesco R, Migliaccio G, Paonessa G;  
 DR MPI; 2002-599793/64.  
 XX  
 PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and  
 PT expression.  
 PS Claim 9; Page; 69pp; English.

The invention relates to nucleic acid molecules comprising altered HCV  
 NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 internal ribosome entry site (IRES) region coding for one or more NS3,  
 NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 are detailed in the specification. Also included are (1) an expression  
 vector comprising a nucleotide sequence coding for the altered nucleic  
 acid, which is transcriptionally coupled to an exogenous promoter; (2) a  
 recombinant cell human hepatoma cell comprising the altered nucleic acids  
 ; (3) a recombinant cell produced by introducing into a human hepatoma  
 cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 replicon enhanced cell or which containing a functional HCV replicon; (5)  
 an HCV replicon enhanced cells made in the method; and (6) measuring the  
 ability of a compound to affect HCV activity. The HCV replicons and HCV  
 expression, and HCV and host cell interactions, producing HCV RNA and  
 proteins, and providing a system for measuring the ability of a compound  
 to modulate one or more HCV activities e.g. to discover drugs which may

CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
 CC mutant of the invention. Note: The present sequence is not shown in the  
 CC specification but was created by the indexer using the HCV sequence  
 CC appearing as ABR91411 and the information in Claim 9  
 XX  
 SQ Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 U; 0 Other;

Query Match 86.9%; Score 1572.2; DB 6; Length 9605;  
 Best Local Similarity 91.8%; Pred. No. 0;  
 Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1 ATGAGCAGATCTCTAAACCTCAAGAAACCAACGTAACACCAACCGCCGACAG 60  
 DB 342 ATGAGCAGATCTCTAAACCTCAAGAAACCAACGTAACACCAACCGCCGACAG 401  
 QY 61 GAGCTCAAGTCCCGGCGGCTGATGATGCTTGTGAGTTTACCTGTGCGCGAG 120  
 DB 402 GAGCTCAAGTCCCGGCGGCTGATGATGCTTGTGAGTTTACCTGTGCGCGAG 461  
 QY 121 GGGCCCAAGTTGGGTGTGCGCGCACTAGAGAACTTCCGACCGGTCCGAACCTGTGA 180  
 DB 462 GGGCCCAAGTTGGGTGTGCGCGCACTAGAGAACTTCCGACCGGTCCGAACCTGTGA 521  
 QY 181 AGGCGACACTATCCCAAGGCTGCGCAGCCGAGGCGGCTGAGCCGCG 240  
 DB 522 AGGCGACACTATCCCAAGGCTGCGCAGCCGAGGCGGCTGAGCCGCG 581  
 QY 241 TACCTTGGCCCTCTATAGCAAGAGGCAATGGGGTGGGAGAGATGCTCTGTACCC 300  
 DB 582 TACCTTGGCCCTCTATAGCAATGGGGTGGGAGAGATGCTCTGTACCC 641  
 QY 301 CGTGGCTCCCGGCTATGTTGGGGCCCACTGACCCCGGCTATAGCTTGTGGGT 360  
 DB 642 CGTGGCTCCCGGCTATGTTGGGGCCCACTGACCCCGGCTATAGCTTGTGGGT 701  
 QY 361 AAAGTCATGATACCTTCATGCGGCTTGGCGACCTATGAGGAGTACCTGCTGCTG 420  
 DB 702 AAAGTCATGATACCTTCATGCGGCTTGGCGACCTATGAGGAGTACCTGCTGCTG 761  
 QY 421 GGGCTCCCTTATAGGGGCGCTGCAAGGCGCTGAGGCAATGGGCTTGTGAGAG 480  
 DB 762 GGGCTCCCTTATAGGGGCGCTGCAAGGCGCTGAGGCAATGGGCTTGTGAGAG 821  
 QY 481 GGGCTGAATATAGCAAGGAAATTTACCGGCTGCTTCTATCTCTCTGCTGGCT 540  
 DB 822 GGGCTGAATATAGCAAGGAAATTTACCGGCTGCTTCTATCTCTCTGCTGGCT 881  
 QY 541 TTGCTGCTGTTTGAACAATCCAGCTTCCGCTTATGAGTGGGAGTCCGGGATC 600  
 DB 882 TTGCTGCTGTTTGAACAATCCAGCTTCCGCTTATGAGTGGGAGTCCGGGATC 941  
 QY 601 TACCATGTACGAGAGATGCTCAACTCAAGCATCTGTAGAGACAGCGGACATGATC 660  
 DB 942 TACCATGTACGAGAGATGCTCAACTCAAGCATCTGTAGAGAGAGCGGACATGATC 1001  
 QY 661 ATGACACCCCGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 DB 1002 ATGACACCCCGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1061  
 QY 721 GGGCTCACTCCAGCGTGGCGGCAAGAGCGGAGAGATCCCACTGCGCAATTAAGAGCG 780  
 DB 1062 GGGCTCACTCCAGCGTGGCGGCAAGAGCGGAGAGATCCCACTGCGCAATTAAGAGCG 1121  
 QY 781 CAGCTGATTTGCTGTTGGGCGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 DB 1122 CAGCTGATTTGCTGTTGGGCGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181  
 QY 841 TGGGATGCTGTTTCTGCTCTCAAGCTGTTCACTTCCGCTCCGCAATCAATCAAGAG 900  
 DB 1182 TGGGATGCTGTTTCTGCTCTCAAGCTGTTCACTTCCGCTCCGCAATCAATCAAGAG 1241  
 QY 901 GTACAGGACTGCAATTTGTTCAATCTATCCGGGCAAGTATAGGTACCGCATGCTTGG 960









[illegible]

Query	Match	Best Local Similarity	Score	Pred. No.	DB	Length	Sequence
DB	1661: Conservative	0	Mismatches	148	Indels	0	Gaps
QY	1	ATGAGCAGCAATCTTAAACCTCAAGAAAAACAAAGTAACACCAACCGCGCCACAG	60				
DB	342	ATGAGCAGCAATCTTAAACCTCAAGAAAAACAAAGTAACACCAACCGCGCCACAG	401				
QY	61	GACGTCAGATTCGCGGGCGGTGTCAGATCGTGTGTGAGATTACCTGTTCGCGCAGG	120				
DB	402	GACGTCAGATTCGCGGGCGGTGTCAGATCGTGTGTGAGATTACCTGTTCGCGCAGG	461				
QY	121	GGCCCCAGGTGGGTGTGCGCGGCACCTAGAAAGCTTCCAGCGGTGGCAACCTCGTGGG	180				
DB	462	GGCCCCAGGTGGGTGTGCGCGGCACCTAGAAAGCTTCCAGCGGTGGCAACCTCGTGGG	521				
QY	181	AGGCGACAACTTATCCCAAGGCTCGGCAGCGCCAGAGGCGTGGGCTCAAGCCCGAG	240				
DB	522	AGGCGACAACTTATCCCAAGGCTCGGCAGCGCCAGAGGCGTGGGCTCAAGCCCGAG	581				
QY	241	TACCTTGTGCCCCCTCTATAGGCAAGAGGGCATGGGGTGGCGAGATGGCTCTGTACCC	300				
DB	582	TACCTTGTGCCCCCTCTATAGGCAAGAGGGCATGGGGTGGCGAGATGGCTCTGTACCC	641				
QY	301	CGTGGCTCCGGCCTAGTTGGGGGCCCACTGACCCCGGGCTAGTGGCGGTAATTTGGGT	360				
DB	642	CGTGGCTCCGGCCTAGTTGGGGGCCCACTGACCCCGGGCTAGTGGCGGTAATTTGGGT	701				
QY	361	AAAGTCATCGATCACCTTCACATATGGGGCTCGCGACCTCATGGGGTCAATTCCGCTCGT	420				
DB	702	AAAGTCATCGATCACCTTCACATATGGGGCTCGCGACCTCATGGGGTCAATTCCGCTCGT	761				
QY	421	GGCGCTCCCTTATGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCGTCCGGGTTCTGAGAGC	480				
DB	762	GGCGCTCCCTTATGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCGTCCGGGTTCTGAGAGC	821				
QY	481	GGCGTGAATATGCAACAGGGAAATTACCCGGGTGTCTTCTCTATCTTCTCTTGGCT	540				
DB	822	GGCGTGAATATGCAACAGGGAAATTACCCGGGTGTCTTCTCTATCTTCTCTTGGCT	881				
QY	541	TTGCTGTCCTTATGACCATTCACGATTCGCTTATGAAAGTGGCGCAACGTATCCGGAGT	600				
DB	882	TTGCTGTCCTTATGACCATTCACGATTCGCTTATGAAAGTGGCGCAACGTATCCGGAGT	941				
QY	601	TACCATGTCAAGACGATTCCTCAACTCAAGCATCTGTATCGAGACAGCGGACATGATC	660				
DB	942	TACCATGTCAAGACGATTCCTCAACTCAAGCATCTGTGTATGAGGACAGCGGACATGATC	1001				
QY	661	ATGCAACCCCCGGGTGTGCTCTGTGTCCGGAGGGTAAATTCCTCCGCTGTGGGTA	720				

Db 1002 ATGCATACCCCGGGGTCGTGCCCTGCTGGGAGAACAACTCTCCCGCTGCGGAT 1061  
 QY 721 GGCTCACTCCCAAGCTCGCGGCAAGAGAGCGCAGAGATCCCACTGCGCAATTAAGAGCG 780  
 Db 1062 GGGCTACTCCCAAGCTCGCGGCAAGAGAGCGCTAGGCTCCCACTAGCAATTAAGAGCG 1121  
 QY 781 CAGCTGATTTGCTCGTTGGGGGCGGCTTCCTGCTCCGCTAATGATCGTGGGAGATCTC 840  
 Db 1122 CATGTCGATTTGCTCGTTGGGGGCGGCTTCCTGCTCCGCTAATGATCGTGGGAGATCTC 1181  
 QY 841 TGGCGATCTGTTTCT 900  
 Db 1182 TGGCGATCTGTTTCT 1241  
 QY 901 GTACAGGATCTGCAATGTTCAATCTATCCCGGCGCAAGTACAGTACCGAGCTGG 960  
 Db 1242 GTACAGGATCTGCAATGTTCAATCTATCCCGGCGCAAGTACAGTACCGAGCTGG 1301  
 QY 961 GATATGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 Db 1302 GATATGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361  
 QY 1021 CCACAAAGCTGTGCTGAGCAATGCTGAGCGGCGGCGCACTGGGAGTCTGGCGGCTTGGC 1080  
 Db 1362 CCACAAAGCTGTGCTGAGCAATGCTGAGCGGCGGCGCACTGGGAGTCTGGCGGCTTGGC 1421  
 QY 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
 Db 1422 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1481  
 QY 1141 GTTGAACGGGGAACCTTTACAGACAGAGGGGAGACAGAGCGCGCGCCACAGCGGCTTACA 1200  
 Db 1482 GTTGAACGGGGAACCTTTACAGACAGAGGGGAGACAGAGCGCGCGCCACAGCGGCTTACA 1541  
 QY 1201 TCCCTCTTACACCTGCGGCGGCGGCTGAGAAATTCAGCTGTTAAACCAACGCGCACTG 1260  
 Db 1542 TCCCTCTTACACCTGCGGCGGCGGCTGAGAAATTCAGCTGTTAAACCAACGCGCACTG 1601  
 QY 1261 CACATCAACAGAACTGCTTGAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 Db 1602 CACATCAACAGAACTGCTTGAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1661  
 QY 1321 CTGTTCTACAGCAAGTTCAATGCTGCGAGTCTGAGCGGCGCAAGTCCGCGC 1380  
 Db 1662 CTGTTCTACAGCAAGTTCAATGCTGCGAGTCTGAGCGGCGCAAGTCCGCGC 1721  
 QY 1381 CCCATTGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
 Db 1722 CCCATTGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1781  
 QY 1441 GACCAAGAGGCGCTAATGCTGAGCACTGACCTCAACCGTGTGATCGTGC CGCGGCTTG 1500  
 Db 1782 GACCAAGAGGCGCTAATGCTGAGCACTGACCTCAACCGTGTGATCGTGC CGCGGCTTG 1841  
 QY 1501 CAGGTGTGTCGCCAGTGTACTGTTCACTCCAGGCCCTGTTGTGTTGGGAGCAACCAT 1560  
 Db 1842 CAGGTGTGTCGCCAGTGTACTGTTCACTCCAGGCCCTGTTGTGTTGGGAGCAACCAT 1901  
 QY 1561 CGTTTTCGGGCGGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 Db 1902 CGTTTTCGGGCGGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1961  
 QY 1621 AACACGCGGCGCGCACAGGGGCAACTGTGCGCTGATCACTGATGATGATGATGATGATGAT 1680  
 Db 1962 AACACGCGGCGCGCACAGGGGCAACTGTGCGCTGATCACTGATGATGATGATGATGATGAT 2021  
 QY 1681 AACACGCGGCGCGCACAGGGGCGCGCGTGCACATCGGGGCGGCTGCGCAACAACCTTGTATC 1740  
 Db 2022 AACACGCGGCGCGCACAGGGGCGCGCGTGCATATCGGGGCGGATGCGGCAATTAACCTTGAAC 2081  
 QY 1741 TACCCCAAGGATGCTGCTTCCGGAAGCATCCGAGGCGCACTTACACCAATGCGGTTGCGGG 1800  
 Db 2082 TACCCCAAGGATGCTGCTTCCGGAAGCATCCGAGGCGCACTTACACCAATGCGGTTGCGGG 2141

QY 1801 CTTGTGTTG 1809  
 Db 2142 CTTGTGTTG 2150  
 RESULT 9  
 ABR91425  
 ID ABR91425 standard; DNA; 9605 BP.  
 XX  
 AC ABR91425;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE Hepatitis C virus Con 1 isolate DNA mutant 2.  
 XX  
 KW HCV, db; Con 1; adaptive mutation; liver failure; cirrhosis;  
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KW internal ribosome entry site; IRBS; NS5A; HCV replication; mutant.  
 OS Hepatitis C virus.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH 342. .9374  
 FT CDS  
 FT /tag= a  
 FT /product= "HCV polypeptide"  
 FT /note= "The polypeptide consists of the Core, E1, E2, P7,  
 FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"  
 FT replace(3946..A)  
 FT mutation  
 FT /\*tag= b  
 PN  
 PD W0200259321-A2.  
 XX  
 XX 01-AUG-2002.  
 XX  
 XX 16-JAN-2002; 2002MO-EP000526.  
 XX  
 XX 23-JAN-2001; 2001US-0263479P.  
 XX  
 XX (RICE-) 1ST RICE RCHB BIOL MOLECULAR ANELETTI.  
 XX  
 XX De Francesco R, Migliaccio G, Paonessa G;  
 PI  
 DR WPI; 2002-599793/64.  
 XX  
 PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRBS) region, useful in studying HCV replication and  
 PT expression.  
 XX  
 XX  
 PS Claim 9; Page; 69pp; English.  
 CC  
 CC The invention relates to nucleic acid molecules comprising altered HCV  
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRBS) region coding for one or more NS3,  
 CC NS5A, or EMCV IRBS mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC expression enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1  
 CC mutant of the invention. Note: The present sequence is not shown in the

CC specification but was created by the indexer using the HCV sequence  
CC appearing as ABK91411 and the information in Claim 9

XX Sequence 9605 BP; 1909 A; 2883 C; 2734 G; 2079 T; 0 U; 0 Other;

Query Match 86.9%; Score 1572.2; DB 6; Length 9605;

Best Local Similarity 91.8%; Pred. No. 0;

Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

```

QY 1 ATGAGCAGAAATCTTAAACCTCAAGAAAAAACAACGTAACACCAACCGCGCCACAG 60
DB 342 ATGAGCAGAAATCTTAAACCTCAAGAAAAAACAACGTAACACCAACCGCGCCACAG 401
QY 61 GAGGTCAATTTCCCGCGCGGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
DB 402 GAGGTCAATTTCCCGCGCGGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 461
QY 121 GAGCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 462 GAGCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 521
QY 181 AGGCGAACAACCTATCCCAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
DB 522 AGGCGAACAACCTATCCCAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 581
QY 241 TACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
DB 582 TACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 641
QY 301 CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
DB 642 CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 701
QY 361 AAGTCATGATGATCCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 702 AAGTCATGATGATCCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 761
QY 421 GGGGTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB 762 GGGGTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 821
QY 481 GGGGTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
DB 822 GGGGTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 881
QY 541 TTTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
DB 882 TTTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 941
QY 601 TACCATGTCAAGAACGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
DB 942 TACCATGTCAAGAACGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1001
QY 661 ATGCAACACCCCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
DB 1002 ATGCAACACCCCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1061
QY 721 GGGGTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 1062 GGGGTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1121
QY 781 CAGGTGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 1122 CAGGTGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1181
QY 841 TGGGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
DB 1182 TGGGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1241
QY 901 GTACAGAGCTGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
DB 1242 GTACAGAGCTGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1301
```

```

QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1302 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
QY 1021 CCACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
DB 1362 CCACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1421
QY 1081 TACTATTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
DB 1422 TACTATTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1481
QY 1141 GTTGAAGGGGAACCTTACAGACAGAGGGGGAACAGAGGCGCGCCACGAGCTTAA 1200
DB 1482 GTTGAAGGGGAACCTTATGTACAGAGGGGGAACAGAGGCGCGCCACGAGCTTAA 1541
QY 1201 TCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
DB 1542 TCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1601
QY 1261 CACATTAACAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
DB 1602 CACATTAACAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1661
QY 1321 CTGTGTACAGCAGAGTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
DB 1662 CTGTGTACAGCAGAGTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1721
QY 1381 CCCATTGACAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
DB 1722 CCCATTGACAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1781
QY 1441 GACCAAGGCGCTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
DB 1782 GACCAAGGCGCTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1841
QY 1501 CAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
DB 1842 CAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1901
QY 1561 CGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
DB 1902 CGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1961
QY 1621 AACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1680
DB 1962 AACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2021
QY 1681 ACCAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
DB 2022 ACCAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2081
QY 1741 TGCCCGACGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
DB 2082 TGCCCGACGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2141
QY 1801 CCTTGTGTG 1809
DB 2142 CCTTGTGTG 2150

RESULT 10
ABK91426
ID ABK91426 standard; DNA; 9605 BP.
XX
AC ABK91426;
XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate DNA mutant 3.
XX
KM HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
```







```
Db      1542 TCCTCTTTTACCCGGGATCCAGAAATCCAGTTGTAACCAAGGACGTGG 1601
Qy      1261 CACATCAACGAAGCTGCTTGAAGTGAATGACTCCCTCCAACTGGGTTCTTGCCGC 1320
Db      1602 CACATCAACGAAGCTGCTTGAAGTGAATGACTCCCTCCAACTGGGTTCTTGCCGC 1661
Qy      1321 CTTGTTACACGACGAGTTTCATGCGTCCGAGTGCATGAGGAGGACCTGCGCG 1380
Db      1662 CTTGTTACGAGTCAAGTTCAATCTCATCTGAGATGCCAGAGGCGCATGGCTGAGC 1721
Qy      1381 CCCATTGACCAAGTTTCATGAGGAGTGGGTCCTCATCTTATATAGTCCCAAGCTTG 1440
Db      1722 CCCATTGACCAAGTTTCATGAGGAGTGGGTCCTCATCTTATATAGTCCCAAGCTTG 1781
Qy      1441 GACCAAGGCGCTTATGCTGAGCACTACGACCTCAACCGTGTGTATCGTCCCGGTTG 1500
Db      1782 GACCAAGGCGCTTATGCTGAGCACTACGACCTCAACCGTGTGTATCGTCCCGGTTG 1841
Qy      1501 CAGGTGTGTGGCCAGGTATCTGTTTCACTCCAAAGCCCTGTGTGTGTGGGAGAGCCGAT 1560
Db      1842 CAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1901
Qy      1561 CGTTTGGGCGCCCTTACGTATGATGAGTGTGAGATGAGAGGACGTGCTTCTTCAAC 1620
Db      1902 CGGTGTGGGCGCTTACGTATGATGAGTGTGAGATGAGAGGACGTGCTTCTTCAAC 1961
Qy      1621 AACACGCGGCGCCGACGAGGCACTGTGCTGTATGATGATGATGATGATGATGATGATG 1680
Db      1962 AACACGCGGCGCCGACGAGGCACTGTGCTGTATGATGATGATGATGATGATGATGATG 2021
Qy      1681 ACBAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
Db      2022 ACBAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2081
Qy      1741 TGCCCAACGAGCTGCTTCCGGAAGATCCCGAGGCACTTACCAAAATGCGGTGCGAG 1800
Db      2082 TGCCCAACGAGCTGCTTCCGGAAGATCCCGAGGCACTTACCAAAATGCGGTGCGAG 2141
Qy      1801 CCTGTGTTG 1809
Db      2142 CCTGTGTTG 2150

RESULT 11
ABK91433
ID      ABK91433 standard; DNA; 9605 BP.
AC      ABK91433;
XX
DT      15-NOV-2002. (first entry)
XX
DE      Hepatitis C virus Con 1 isolate DNA mutant 10.
XX
KM      HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KM      hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KM      internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
XX
OS      Hepatitis C virus.
OS      Synthetic.
XX
XX      Key      Location/Qualifiers
FT      CDS      342..9374
FT      /tag= a
FT      /product= "HCV polyprotein"
FT      /note= "The polyprotein consists of the Core, E1, E2, P7,
FT      NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT      mutation
FT      replace(6953,C)
FT      /*tag= b
XX
XX      MO200259321-A2.
XX      01-AUG-2002.
XX
```

```
PF      16-JAN-2002; 2002MO-EP000526.
XX
XX      23-JAN-2001; 2001JUS-0263479P.
XX
XX      (RICE-) 1ST RICEGRHE BIOL MOLECULAR ANGELLETTI.
XX
PI      De Francesco R, Migliaccio G, Paonessa G;
XX      WPI; 2002-599793/64.
XX
XX      New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX      NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX      ribosome entry site (IRES) region, useful in studying HCV replication and
XX      expression.
XX
XX      Claim 9; Page; 69pp; English.
XX
XX      The invention relates to nucleic acid molecules comprising altered HCV
XX      NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV),
XX      internal ribosome entry site (IRES) region coding for one or more NS3,
XX      NS5A, or EMCV IRES mutations, respectively. The location of the mutations
XX      are detailed in the specification. Also included are (1) an expression
XX      vector comprising a nucleotide sequence coding for the altered nucleic
XX      acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX      recombinant cell human hepatoma cell comprising the altered nucleic acids
XX      ; (3) a recombinant cell produced by introducing into a human hepatoma
XX      cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX      replicon enhanced cell or which containing a functional HCV replicon; (5)
XX      an HCV replicon enhanced cells made in the method; and (6) measuring the
XX      ability of a compound to affect HCV activity. The HCV replicons and HCV
XX      replicon enhanced cells are useful in studying HCV replication and
XX      expression, and HCV and host cell interactions, producing HCV RNA and
XX      proteins, and providing a system for measuring the ability of a compound
XX      to modulate one or more HCV activities e.g. to discover drugs which may
XX      treat HCV mediated diseases such as liver failure, cirrhosis and
XX      hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
XX      mutant of the invention. Note: The present sequence is not shown in the
XX      specification but was created by the indexer using the HCV sequence
XX      appearing as ABK91411 and the information in Claim 9
XX
SQ      Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2079 T; 0 U; 1 Other;
XX
Query Match      86.9%; Score 1572.2; DB 6; Length 9605;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
Qy      1 ATGAGCAGCAATCTCTAAACCTCAAGAAAAACCAAGCTATACCAACGCGCCGACAG 60
Db      342 ATGAGCAGCAATCTCTAAACCTCAAGAAAAACCAAGCTATACCAACGCGCCGACAG 401
Qy      61 GACGTCAAGTTCCCGGCGGTGTGTCAGATCGTTGTGAGATTACCTGTTGCCGCGAG 120
Db      402 GACGTCAAGTTCCCGGCGGTGTGTCAGATCGTTGTGAGATTACCTGTTGCCGCGAG 461
Qy      121 GGGCCAGGTTGGGTGTGCGCGGACTAGAAAGACTTCGAGGCGTGCACCTGTGGA 180
Db      462 GGGCCAGGTTGGGTGTGCGCGGACTAGAAAGACTTCGAGGCGTGCACCTGTGGA 521
Qy      181 AGGCGACAACCTATCCCAAGGCTCGCAGCCGAGGCGAGGCGCTCGGCTCAGCCCGG 240
Db      522 AGGCGACAACCTATCCCAAGGCTCGCAGCCGAGGCGAGGCGCTCGGCTCAGCCCGG 581
Qy      241 TACCTTGGCCCTCTATAGGCAACGAGGCGATGAGGTGAGAGATGCTCTGTCAACC 300
Db      582 TACCTTGGCCCTCTATAGGCAACGAGGCGATGAGGTGAGAGATGCTCTGTCAACC 641
Qy      301 CGTGGCTCCCGGCTAGTGGGCGCCCACTGACCCCGCGCGTATGAGGTAATTTGGGT 360
Db      642 CGTGGCTCCCGGCTAGTGGGCGCCCACTGACCCCGCGCGTATGAGGTAATTTGGGT 701
Qy      361 AAAGTCATGATACCTCTCATATGCGGCTTGCAGCACTCATGAGGAGTACATTCGCTGTC 420
Db      702 AAAGTCATGATACCTCTCATATGCGGCTTGCAGCACTCATGAGGAGTACATTCGCTGTC 761
```

```
QY 421 GGCCTCCCTTAGAGGGGCGCTGCAGAGGCGCTGCGCATGAGCGTCCGGGTTCTGAGAGAC 480
Db 762 GGGGCCCCCTTAGGGGGGCGCTGCAGAGGCGCTGCGCATGAGCGTCCGGGTTCTGAGAGAC 821
QY 481 GGGGTGAATCTATGCAACAGGGAAATTTACCCGGTGTCTTTCTCTATCTCTTCTTGGCT 540
Db 822 GGGGTGAATCTATGCAACAGGGAAATCTGCGGGTGTCTTTCTCTCTTCTTCTTGGCT 881
QY 541 TTGCTGTCTGTATGCAATTCAGCTTCGGCTTATGAAGTGGCAAGCTGTCGGGATC 600
Db 882 TTGCTGTCTGTATGCAATTCAGCTTCGGCTTATGAAGTGGCAAGCTGTCGGGATC 941
QY 601 TACCATGTCAAGACGATGCTCACTCAAGCATGCTAGAGACAGAGCATGATC 660
Db 942 TACCATGTCAAGACGATGCTCACTCAAGCATGCTAGAGACAGAGCATGATC 1001
QY 661 ATGCACACCCCGGGGTGTGCTGTGTCCGGAGGGTAATTCCTCCGCTGTGGGTA 720
Db 1002 ATGCATATCCCGGGGTGTGCTGTGTCCGGAGAGAACATCTCCGCTGTGGGTA 1061
QY 721 GGGCTACCTCCAGCGTGGGGCCAGAGAGCGCAGATCCCACTGCGCAATATGAGCGC 780
Db 1062 GGGCTACCTCCAGCGTGGGGCCAGAGAGCGCAGATCCCACTGCGCAATATGAGCGC 1121
QY 781 CAGTCGATTTTCTGCTGGGGCGGCTGCTTCTGCTCCGTAATGATGAGGGGATCTC 840
Db 1122 CAGTCGATTTTCTGCTGGGGCGGCTGCTTCTGCTCCGTAATGATGAGGGGATCTC 1181
QY 841 TGGCGATCTGTTTTCTGCTGCTCTCACTGCTTCACTTCTGCTGCGCGCATCAAGC 900
Db 1182 TGGCGATCTGTTTTCTGCTGCTCTCACTGCTTCACTTCTGCTGCGCGCATCAAGC 1241
QY 901 GTACAGAGCTGCAATTTGTTCAATCTATCCGGGCGAGTATCAGGTACCGGATGCTGG 960
Db 1242 GTACAGAGCTGCAATTTGTTCAATCTATCCGGGCGAGTATCAGGTACCGGATGCTGG 1301
QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 1302 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1361
QY 1021 CCACAAAGCTGTGTGAGCATGTGTGGGGGGGCGCCACTGGGGAGTCTGGCGGCTTGC 1080
Db 1362 CCACAAAGCTGTGTGAGCATGTGTGGGGGGGCGCCACTGGGGAGTCTGGCGGCTTGC 1421
QY 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1422 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1481
QY 1141 GTTGAAGGGGAACTTTACAGACAGAGGGGAGACAGCGCGCGCGCCACCGGCTTACA 1200
Db 1482 GTTGAAGGGGAACTTTACAGACAGAGGGGAGACAGCGCGCGCGCCACCGGCTTACA 1541
QY 1201 TCCCTTTTCAACCTGGGGCGGCTCAAGAAATTCAGCTTTTAAACCAACAGCGAGCTGG 1260
Db 1542 TCCCTTTTCAACCTGGGGCGGCTCAAGAAATTCAGCTTTTAAACCAACAGCGAGCTGG 1601
QY 1261 CACATCAACAGAACTGCTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1602 CACATCAACAGAACTGCTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATG 1661
QY 1321 CTGTTCTACAGCAGGTTTCAATGCTCCGATGCTCAAGCGGATGCGCAGCTGCGC 1380
Db 1662 CTGTTCTACAGCAGGTTTCAATGCTCCGATGCTCAAGCGGATGCGCAGCTGCGC 1721
QY 1381 CCCATTTGACAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Db 1722 CCCATTTGACAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1781
QY 1441 GACGAGAGGCGCTATTTGCTGAGCTACGACCTCAACCGTGTGATGATGATGATGATGATG 1500
Db 1782 GACGAGAGGCGCTATTTGCTGAGCTACGACCTCCCGGCGTGTGATGATGATGATGATGATG 1841
```

```
QY 1501 CAGGTGTGTGGCCAGATGATGTTTCACTCCAAAGCCCTGTGTGTGTGGAGACGACGAT 1560
Db 1842 CAGGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1901
QY 1561 CGTTTGGGCGCCCTTACGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db 1902 CGTTTGGGCGCCCTTACGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1961
QY 1621 AACACGAGGCGCCAGCAGGAGGCAATGTTTGGGCTGATCATGATGATGATGATGATGATG 1680
Db 1962 AACACGAGGCGCCAGCAGGAGGCAATGTTTGGGCTGATCATGATGATGATGATGATGATG 2021
QY 1681 ACCAAGAGTGTGGGGGCGCCCGTCAACATCGGGGGGTGCGCAACACATTTGATC 1740
Db 2022 ACCAAGAGTGTGGGGGCGCCCGTCAACATCGGGGGGTGCGCAACATTTGATC 2081
QY 1741 TGGCCACGAGCTGTTTCCGAAAGATTCGAGGCGACTTACACCAATGCGGTTTGGGG 1800
Db 2082 TGGCCACGAGCTGTTTCCGAAAGATTCGAGGCGACTTACACCAATGCGGTTTGGGG 2141
QY 1801 CTTGGTTG 1809
Db 2142 CTTGGTTG 2150
```

```
RESULT 12
AAD25332
ID AAD25332 standard; cDNA; 9605 BP.
XX
XX AAD25332;
XX
XX 12-MAR-2002 (first entry)
XX
XX Hepatitis C virus (HCV) full-length cDNA mutant #2.
XX
XX DE Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
XX KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
XX KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
XX KW mutant; ss.
XX
XX OS Hepatitis C virus.
XX
XX OS Synthetic.
XX
XX WO200189364-A2.
XX
XX 29-NOV-2001.
XX
XX 23-MAY-2001; 2001WO-US016822.
XX
XX 23-MAY-2000; 2000US-00576989.
XX
XX (UNIV ) UNIV WASHINGTON.
XX
XX PA
XX PI Rice CM, Blight KJ;
XX
XX DR WPI; 2002-066755/09.
XX
XX PT Hepatitis C virus variants having greater transfection efficiency and
XX PT ability to survive subpassage, useful as a vaccine for immunizing primate
XX PT to the virus, comprise non-naturally occurring viral sequences.
XX
XX PS Example 2; Page 97-100; 174pp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
XX polynucleotides comprising non-naturally occurring HCV sequence and HCV
XX variants that have a transfection efficiency and ability to survive
XX subpassage greater than HCV that have wild-type polypeptide coding
XX regions. The polynucleotides of the invention are useful for identifying
XX a cell line that is permissive for infection with HCV and detecting
XX infection of HCV in cells of the cell line. They are also useful for
XX testing a compound for anti-viral properties and for inhibiting HCV
XX infection. They are also useful for the generation of defined HCV virus
XX stocks to develop in vitro and in vivo assays for virus neutralisation,
XX attachment, penetration and entry, structure/function studies on HCV
```

CC proteins and RNA elements and identification of new antiviral targets, a  
CC systematic survey of cell culture systems and conditions to identify  
CC those that support wild-type and variant HCV RNA replication and particle  
CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of attenuated or  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilization of the HCV  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
CC The present sequence is Hepatitis C virus (HCV) full-length cDNA  
CC containing the mutation that results in Ser to Ile at position 1179 of  
CC HCVrepBartman protein

Sequence 9605 BP; 1910 A; 2884 C; 2731 G; 2080 T; 0 U; 0 Other;

Query Match 86.9%; Score 1572.2; DB 6; Length 9605;

Best Local Similarity 91.8%; Pred. No. 0;

Matches 1661; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1 ATGAGCAGCAATCTTAACCTCAAGAAAAACCAACGTACACCAACCGCGCCCAACG 60  
DB 342 ATGAGCAGCAATCTTAACCTCAAGAAAAACCAACGTACACCAACCGCGCCCAACG 401  
QY 61 GAGGTCAAGTTCCTCGGCGCGGTGTCAGAGTGTGTGAGTTCCTGTTGCCGCGAG 120  
DB 402 GAGGTCAAGTTCCTCGGCGCGGTGTCAGAGTGTGTGAGTTCCTGTTGCCGCGAG 461  
QY 121 GGGCCCAAGTTGGGTGTGCGCGCACTAGAGAGACTTCGAGCGGTGCAACTCTGTGA 180  
DB 462 GGGCCCAAGTTGGGTGTGCGCGCACTAGAGAGACTTCGAGCGGTGCAACTCTGTGA 521  
QY 181 AGGCGCAACCTATCCCAAGGCTCGCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
DB 522 AGGCGCAACCTATCCCAAGGCTCGCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 581  
QY 241 TACCTTGGCCCTCTATGAGCAAGAGGATGGGGTGGGAGAGATGGCTCTGTCAACC 300  
DB 582 TACCTTGGCCCTCTATGAGCAAGAGGATGGGGTGGGAGAGATGGCTCTGTCAACC 641  
QY 301 CGTGCTCCGCGCTAGTGGGAGCCCACTGACCCCGCGGTAGTTCGCGTAATTTGGGT 360  
DB 642 CGTGCTCCGCGCTAGTGGGAGCCCACTGACCCCGCGGTAGTTCGCGTAATTTGGGT 701  
QY 361 AAGATATGATATCCTCATATGCGGCTTGCAGACTCATGGGGTATCATTCGCTGCTC 420  
DB 702 AAGATATGATATCCTCATATGCGGCTTGCAGACTCATGGGGTATCATTCGCTGCTC 761  
QY 421 GGGGCTCCCTTAGGGGGGCGCTGCGAGGAGCTGGCGCATGGCGTCCGGGTTCTGAGGAC 480  
DB 762 GGGGCTCCCTTAGGGGGGCGCTGCGAGGAGCTGGCGCATGGCGTCCGGGTTCTGAGGAC 821  
QY 481 GGGGTAACATATGCAACAGGAAATTTACCGGTTGCTTTCTATCTTCTCTTGGCT 540  
DB 822 GGGGTAACATATGCAACAGGAAATTTACCGGTTGCTTTCTATCTTCTCTTGGCT 881  
QY 541 TTGCTGCTCTGTTTGAACATTCAGCTTCGCTTATGAAATGGGCAACGATCCGAGT 600  
DB 882 TTGCTGCTCTGTTTGAACATTCAGCTTCGCTTATGAAATGGGCAACGATCCGAGT 941  
QY 601 TACCATGTACGAAACGATGCTCAACTCAAGCATGCTGTACGAGACAGCGCATGATC 660  
DB 942 TACCATGTACGAAACGATGCTCAACTCAAGCATGCTGTGTATAGGACAGCGCATGATC 1001  
QY 661 ATGCAACCCCGGGGTGTGTCCCTGTGTCGGGAGGATTAATCTCCGCTGCTGGGTA 720  
DB 1002 ATGCAACCCCGGGGTGTGTCCCTGTGTCGGGAGAACAACTCTCCGCTGCTGGGTA 1061

QY 721 GGGCTACCTCCACGCTCGCGCCCAAGAGCAGCAGCATCCCATGTCGAATACGACGC 780  
DB 1062 GGGCTACCTCCACGCTCGCGCCCAAGAGCAGCAGCATCCCATGTCGAATACGACGC 1121  
QY 781 CAGGTGATTTGCTGTTGGGGGCGGTGCTTCTGCTGCTATGTAGCTGGGGATCTC 840  
DB 1122 CAGGTGATTTGCTGTTGGGGGCGGTGCTTCTGCTGCTATGTAGCTGGGGATCTC 1181  
QY 841 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
DB 1182 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241  
QY 901 GTACAGCATGCAATGTTCAATCTATCCCGGCAAGTACAGTACCGCATGCTTGG 960  
DB 1242 GTACAGCATGCAATGTTCAATCTATCCCGGCAAGTACAGTACCGCATGCTTGG 1301  
QY 961 GATATGATGATGAACGTGCTACCTACAGGAGCTGATGATATGCACTCTCCGATC 1020  
DB 1302 GATATGATGATGAACGTGCTACCTACAGGAGCTGATGATATGCACTCTCCGATC 1361  
QY 1021 CCACAGCTGCTGTGACATGCTGGGGGGGCGCACTGGGGAGTCTGGGGGGCTTGGC 1080  
DB 1362 CCACAGCTGCTGTGATATGATGATGGGGGGGCGCACTGGGGAGTCTGGGGGGCTTGGC 1421  
QY 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
DB 1422 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1481  
QY 1141 GTTGAAGGGAACCTTACACGACAGGGGGGACACAGGCGCGCGCCGACGCGCTTACA 1200  
DB 1482 GTTGAAGGGAACCTTACACGACAGGGGGGACAGAGGCGCAAAACACCTCGGGATTAAG 1541  
QY 1201 TCCCTTTTCAACCTGGGGCGGCTCAAGAAATTCAGCTTGTAAACCAACGCGACGTG 1260  
DB 1542 TCCCTTTTCAACCTGGGGCGGCTCAAGAAATTCAGCTTGTAAACCAACGCGACGTG 1601  
QY 1261 CACATCAACAGAACTGCTTGAACGCAATGACCTCCCTCAAACTGGGGTCTTGGCGCG 1320  
DB 1602 CACATCAACAGAACTGCTTGAACGCAATGACCTCCCTCAAACTGGGGTCTTGGCGCG 1661  
QY 1321 CTGTTCTACACGACAGTTCATGCTGCTGATGCTCAGAGCGCATGCGCAGCTGCGC 1380  
DB 1662 CTGTTCTACACGACAGTTCATGCTGCTGATGCTCAGAGCGCATGCGCAGCTGCGC 1721  
QY 1381 CCCATGACCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
DB 1722 CCCATGACCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1781  
QY 1441 GACGAGAGGCGCTATGCTGAGCACTACGCACTCAACCGGTGATGATGATGATGATGATGAT 1500  
DB 1782 GACGAGAGGCGCTATGCTGAGCACTACGCACTCAACCGGTGATGATGATGATGATGATGAT 1841  
QY 1501 CAGGTGTGTGCGCAGTGTACTGTTCACTCCAGACCTGTTGGTGGGAGCAGACCAT 1560  
DB 1842 CAGGTGTGTGCGCAGTGTACTGTTCACTCCAGACCTGTTGGTGGGAGCAGACCAT 1901  
QY 1561 CGTTTGGCGCCCTTACGTAAGATGGGTGAGAAATGAGACGAGCGTGTCTTCAAC 1620  
DB 1902 CGTTTGGCGCCCTTACGTAAGATGGGTGAGAAATGAGACGAGCGTGTCTTCAAC 1961  
QY 1621 AACAGCGGCGCGCAGGGGCAACTGTTCCGCTGATACATGATGATGATGATGATGATGATGAT 1680  
DB 1962 AACAGCGGCGCGCAGGGGCAACTGTTCCGCTGATACATGATGATGATGATGATGATGATGAT 2021  
QY 1681 ACCAAGACGTGTGGGGGCGCCCGGTCACATCGGGGGGTGGGCAACAACCTTGTATC 1740  
DB 2022 ACCAAGACGTGTGGGGGCGCCCGGTCACATCGGGGGGTGGGCAACAACCTTGTATC 2081  
QY 1741 TGCCTCAGGACTGCTTCCGGAAGATCCGAGGCCACTTACACCAATGCGGTTGGGG 1800  
DB 2082 TGCCTCAGGACTGCTTCCGGAAGATCCGAGGCCACTTACACCAATGCGGTTGGGG 2141  
QY 1801 CCTTGCTG 1809





OY	181	AGCGGCAACAAC	TATATCCCAAGGCTGCCAGCCCGAGGGGCAAGGCGCTGGGCTCAGCCCGGG	240
Db	1979	AGCGCAACAAC	CTATCCCAAGGCTGCCAGCCCGAGGGGCTGGGCTCAGCCCGGG	2038
OY	241	TACCCCTTGGCCCT	CTATATGGCAACAAGGAGATGGGGTGGGCAAGATGGCTCCTGTCAACC	300
Db	2039	TACCCCTTGGCCCT	CTATATGGCAATAGGGCTTGGGGTGGGCAAGATGGCTCCTGTCAACC	2099
OY	301	CGTGGCTCCGGCCT	AGTTGGGGGCCCACTGACCCCCGGCGTGGTCCGCTAATTTGGGT	360
Db	2099	CGTGGCTCTGGGCT	AGTTGGGGGCCCACTGACCCCCGGCGTGGTCCGCTAATTTGGGT	2158
OY	361	AAAGTCATCCGAT	CAACCTTCAATGGCGGCTTTCGGCCACTCATGGGGTATCAATTCGCTGCTC	420
Db	2159	AAGTCATCCGAT	CAACCTTCAACCTTCAACGTCGGCTTTCGGCCACTCATGGGGTATCAATTCGCTGCTC	2218
OY	421	GGCGCTCCCTT	AGGGGGGGCGCTGCCAGGGGCCCTGGGCGCATGGCGTCCGGGTTCTGGAGAC	480
Db	2219	GGCGCCCCCTT	AGGGGGGGCGCTGCCAGGGGCCCTGGGCGCATGGCGTCCGGGTTCTGGAGAC	2278
OY	481	GGCGTGAAC	CTATGCAACAGGGAAATTTACCCGGTCTCTTTCTATCTTCTCTTGGCT	540
Db	2279	GGCGTGAAC	CTATGCAACAGGGAAATTTACCCGGTCTCTTTCTATCTTCTCTTGGCT	2338
OY	541	TTGCTGTCTCT	TTTGAACAATTCACACTTCGCTTATGAAATGCGCAAGTGTCCGGATC	600
Db	2339	TTGCTGTCTCT	TTTGAACAATTCACACTTCGCTTATGAAATGCGCAAGTGTTCGGAGTg	2398
OY	601	TACCATGT	CAGGAAGATTGCTCCAACTCAAGCATCGTACGGAACAGCGGACATGATC	660
Db	2399	TACCATGT	CAGGAAGATTGCTCCAACTCAAGCATGTTGTATGAGGACAGCGGACATGATC	2458
OY	661	ATGCAACACCC	CGGGTGTGTGCTGTGTCCGGAGGGTAAATTCCTCCGCTCGTGGGTA	720
Db	2459	ATGCAATACCC	CGGGTGTGTGCTGTGTCCGGAGGGTAAATTCCTCCGCTCGTGGGTA	2518
OY	721	GCGCTCAT	CCCCGCGTCCAGGACGCGCAGCATCCCATCTGCGACATATGACAGCG	780
Db	2519	GCGCTCAT	CCCCGCGTCCAGGACGCGTCCAGGACGCGTCCCATCTGCGACATATGACAGCG	2578
OY	781	CACGTCGAT	TGTGCTGTGGGGCGGCTGCTGTGCTCCGCTATGATGATGGGGGAGATCTC	840
Db	2579	CATGTGCA	TGTGCTGTGGGGCGGCTGCTGTGCTCCGCTATGATGATGGGAGATCTC	2638
OY	841	TGCGGATCTG	TTTTCTGCTCTCAGCTGTTCACCTTCCGCTCCGCGACATCAGACG	900
Db	2639	TGCGGATCTG	TTTTCTGCTGCGGCCAGACTGTTCACCTTCCGCTCCGCGACAGACGA	2699
OY	901	GTCACGAC	CTGCATTTGTTCAATCTATCCCGGCGCATCATAGTACCGCATGGCTTGG	960
Db	2699	GTCACGAC	CTGCATTTGCTCAATATATCCCGGCGCATCATAGTACCGTATGGCTTGG	2758
OY	961	GATATGAT	GATGCAACTGGTCACTCAACAGGCCCTTATGATATGGCACTACTCCGATC	1020
Db	2759	GATATGAT	GATGCAACTGGTCACTCAACAGGCCCTTATGATATGGCACTTATCTCGGATC	2818
OY	1021	CCACAAGCTG	TGTCGATGTCGCGGGGCCCATCTGGGAGTCTTGGCGGGCCTTGGC	1080
Db	2819	CCACAAGCTG	TGTCGATGATGATGATGATGGCGGGGCCCATTTGGGAGTCTTACGGGGCCTTGGC	2878
OY	1081	TACTATTC	TCATGATGGGAGAACTGGGCTAAGGTCTTGTGTGATGCTACTTTGCCGGC	1140
Db	2879	TACTATTC	TCATGATGGGAGAACTGGGCTAAGGTCTTGTGTGATGATCTACTTTGCCGGC	2938
OY	1141	GTTTGA	CGGGGAACCTTACACGACAGGGGGGACACAGCGCGCGGCCCATGGCTTAC	1200
Db	2939	GTTTGA	CGGGGAACCTTATGTGACAGGGGGGAGATGGCCAAAACACCTCGGGATTTACG	2998
OY	1201	TCCCTCTT	CAACCTGCGGCGGCTCAGAAAATCCAGCTTGTAAACCAACGCGCAGCTGG	1260
Db	2999	TCCCTCTTTT	CACCGGGCTCATCCAGAAAATCCAGCTTGTAAACCAACGCGCAGCTGG	3058
OY	1261	CACATCAAC	AGAACTGCTTGAATCGCAATGCTCCCTTCAAACTGGGTTCTTCCCGCG	1320

[illegible]



XX This invention describes a novel Hepatitis C virus (HCV) cell culture system comprising human hepatoma cells that contain an integrated HCV-RNA construct (I). (I) contains the HCV-specific RNA segments 5'-NTR (non-translated region), NS (non-structural)1,3, NS4A, NS4B, NS5A, NS5B and 3'-NTR, and a selectable (marker) gene (II). The cell cultures, and/or (I), are used to prepare, evaluate and/or test therapeutic and/or diagnostic agents for HCV infections, and to prepare vaccines against HCV infection (particularly preparation of attenuated HCV). The can also be used for preparation of a liver-specific delivery system for gene therapy, and to identify cells permissive for HCV replication. Virus RNA replicates autonomously and with high efficiency in this cellular system, so that variations in replication rates can be measured (for screening antiviral agents) quantitatively or qualitatively, using standard laboratory equipment. Efficient replication of HCV RNA is only achieved when the specified RNA segments are present and when the transfected cells are maintained under permanent selection pressure

Sequence 11076 BP; 2221 A; 3297 C; 3149 G; 2409 T; 0 U; 0 Other;

Query Match 86.8%; Score 1570.6; DB 3; Length 11076;

Best Local Similarity 91.8%; Pred. No. 0;

Matches 1660; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 1 ATGAGCAGCAATCTTAACCTCAAGAAAACCAACGTACCAACCGCGCCACAG 60  
DB 1813 ATGGGACAAATCTTAACCTCAAGAAAACCAACGTACCAACCGCGCCACAG 1872  
QY 61 GAGGTCAAGTCCCGGCGGTGTGAGATGTTGTGTGAGATTACCTGTTGCGCGCAG 120  
DB 1873 GAGGTCAAGTCCCGGCGGTGTGAGATGTTGTGTGAGATTACCTGTTGCGCGCAG 1932  
QY 121 GGGCCAGGTTGGGTGTGCGCGGAGTACGAGACCTTCCGAGCGTGCACCTGTGGA 180  
DB 1933 GGGCCAGGTTGGGTGTGCGCGGAGTACGAGACCTTCCGAGCGTGCACCTGTGGA 1992  
QY 181 AGGCGCAACCTATCCCGCAAGGTCGCCAGCCCGAGGAGGAGCGGCGTCAAGCCCGG 240  
DB 1993 AGGCGCAACCTATCCCGCAAGGTCGCCAGCCCGAGGAGGAGCGGCGTCAAGCCCGG 2052  
QY 241 TACCTTGGGCGCTCTATGAGCAACAGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 300  
DB 2053 TACCTTGGGCGCTCTATGAGCAACAGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 2112  
QY 301 CGTGGCTCCGCGCTAGTGGGCGCCCACTGACCCCGCGGTAGTCCGCTAATTTGGGT 360  
DB 2113 CGTGGCTCCGCGCTAGTGGGCGCCCACTGACCCCGCGGTAGTCCGCTAATTTGGGT 2172  
QY 361 AAGGTATGATATCCTTCAATGCGGCTTCCGCACTCATGAGGATGATATCCTTCCGCTGTC 420  
DB 2173 AAGGTATGATATCCTTCAATGCGGCTTCCGCACTCATGAGGATGATATCCTTCCGCTGTC 2232  
QY 421 GGGCGCTCCCTTAGGGGCGCTGCGAGGCGCTTGGCGCATGAGGCTTGGAGTTCGAGAGAC 480  
DB 2233 GGGCGCTCCCTTAGGGGCGCTGCGAGGCGCTTGGCGCATGAGGCTTGGAGTTCGAGAGAC 2292  
QY 481 GGGCGTACATGACCAAGGAAATTTACCGGTTGCTTTCTCTATCTTCTCTTGGCT 540  
DB 2293 GGGCGTACATGACCAAGGAAATTTACCGGTTGCTTTCTCTATCTTCTCTTGGCT 2352  
QY 541 TTGCTGTCTGTTTGAACATTCAGATTCCGCTTAAGAGTGGGCAACGTGTCCGGGATC 600  
DB 2353 TTGCTGTCTGTTTGAACATTCAGATTCCGCTTAAGAGTGGGCAACGTGTCCGGGATC 2412  
QY 601 TACCATGTCAGCAACATTTGCTCAACTCAAGCATGTTGATGACAGCGGACATGATC 660  
DB 2413 TACCATGTCAGCAACATTTGCTCAACTCAAGCATGTTGATGACAGCGGACATGATC 2472  
QY 661 ATGACACCCCGGGGTGTGTGCTGTGTGTCGGGAGGATTAATTTCTTCCGCTCTGGGTA 720  
DB 2473 ATGACACCCCGGGGTGTGTGCTGTGTGTCGGGAGGATTAATTTCTTCCGCTCTGGGTA 2532  
QY 721 GGGCTCACTCCAGGCTCGGCGGCAAGACGCCAGCATCCCACTGGAGCAATACGACGC 780

DB 2533 GGGCTCACTCCAGGCTCGGCGGCAAGACGTAAGTCCCACTACGATACGACGCC 2592  
QY 781 CAGGTGATTTGCTGTGGGCGGCTGCTTGTGCTGCTGCTATGTACGTGGGATCTC 840  
DB 2593 CAGGTGATTTGCTGTGGGCGGCTGCTTGTGCTGCTGCTATGTACGTGGGATCTC 2652  
QY 841 TGGGATCTGTTTCTGCTCTCTACGCTGTTTCACTTTTCCGCTGCGGACATTCAGAC 900  
DB 2653 TGGGATCTGTTTCTGCTCTCTACGCTGTTTCACTTTTCCGCTGCGGACATTCAGAC 2712  
QY 901 GATACAGATGCAATGTTCAATCTATCCCGGCAAGTATGACAGTACCGATGCTTGG 960  
DB 2713 GATACAGATGCAATGTTCAATCTATCCCGGCAAGTATGACAGTACCGATGCTTGG 2772  
QY 961 GATATGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
DB 2773 GATATGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2832  
QY 1021 CCACAAAGCTGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
DB 2833 CCACAAAGCTGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2892  
QY 1081 TACTATTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
DB 2893 TACTATTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2952  
QY 1141 GTTGAAGGGAACCTTACACGACAGGGGGAACACAGGCGCGGCCCAAGGCTTAC 1200  
DB 2953 GTTGAAGGGAACCTTACATGACAGGGGGAACAGATGCCAAAACCTTCCGAGATTAC 3012  
QY 1201 TCCCTTTTCAACCTGAGGCGGCTCAGAAATTCAGCTTGTAAACCAACGCAAGCTG 1260  
DB 3013 TCCCTTTTCAACCTGAGGCGGCTCAGAAATTCAGCTTGTAAACCAACGCAAGCTG 3072  
QY 1261 CACATCAACAGATCTGTTGAACCTGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
DB 3073 CACATCAACAGATCTGTTGAACCTGATGATGATGATGATGATGATGATGATGATGATGATG 3132  
QY 1321 CTGTTTCAACGACAGGTTCAATGCGTCCGATGCTCAGAGGCAAGGCAAGGCTGCGG 1380  
DB 3133 CTGTTTCAACGACAGGTTCAATGCGTCCGATGCTCAGAGGCAAGGCAAGGCTGCGG 3192  
QY 1381 CCCATTGACCAAGTTCATGAGGAGTGGGATCCCATCATTTATATATGATGCCAGCTTG 1440  
DB 3193 CCCATTGACCAAGTTCATGAGGAGTGGGATCCCATCATTTATATATGATGCCAGCTTG 3252  
QY 1441 GACCAAGGCGCTATGCTGAGCACTACGACCTCAACCGTGTGATGATGATGATGATGATG 1500  
DB 3253 GACCAAGGCGCTATGCTGAGCACTACGACCTCAACCGTGTGATGATGATGATGATGATG 3312  
QY 1501 CAGGTGTGAGCCCAAGTATCTGTTCACTCAAGGCTGTGTGTGGGGAAGACCGAT 1560  
DB 3313 CAGGTGTGAGCCCAAGTATCTGTTCACTCAAGGCTGTGTGTGGGGAAGACCGAT 3372  
QY 1561 CTTTTGGGCGCCCTTACATGATGAGGATGAGATGAGACCGACGTGCTGCTTCAAC 1620  
DB 3373 CTTTTGGGCGCCCTTACATGATGAGGATGAGATGAGACCGACGTGCTGCTTCAAC 3432  
QY 1621 AACAGCGGCGCGCAAGGGAACCTGTTCCGCTGATGATGATGATGATGATGATGATGATG 1680  
DB 3433 AACAGCGGCGCGCAAGGGAACCTGTTCCGCTGATGATGATGATGATGATGATGATGATG 3492  
QY 1681 ACCAAGCATGTGGGGGCGCCCGTGTCAACATCGGGGGGTGCGGCAACACTTTGATC 1740  
DB 3493 ACCAAGCATGTGGGGGCGCCCGTGTCAACATCGGGGGGTGCGGCAACACTTTGATC 3552  
QY 1741 TGGCCCAAGGATGCTTCCGGAAGCATCCCGAGGCGCACTTAACCAAAATGCGTTGGGG 1800  
DB 3553 TGGCCCAAGGATGCTTCCGGAAGCATCCCGAGGCGCACTTAACCAAAATGCGTTGGGG 3612  
QY 1801 CTTGGTTG 1809



Db 3613 CCTGCTG 3621

Search completed: February 19, 2005, 01:26:50  
Job time : 1018.71 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2005, 00:21:46 ; Search time 311.302 Seconds  
(without alignments)  
9508.540 Million cell updates/sec

Title: US-09-664-363-21\_COPY\_308\_2116

Perfect score: 1809

Sequence: 1 ATGACGACGAATCTTAAC.....GGGTTGGGGCTTGTGTTG 1809

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1809	100.0	2116	3	US-08-191-160-21
2	1570.6	86.8	11076	4	US-09-539-601-1
3	1570.6	86.8	11076	4	US-09-539-601-19
4	1570.6	86.8	11076	4	US-09-539-601-25
5	1570.6	86.8	11076	4	US-09-539-601-31
6	1556.2	86.0	9595	3	US-09-014-416-4
7	1556.2	86.0	9595	3	US-09-014-416-6
8	1543.4	85.3	9413	4	US-09-827-688-6
9	1538.6	85.1	2433	3	US-08-612-973-49
10	1538.6	85.1	2433	3	US-08-612-973-49
11	1531.2	84.6	9472	4	US-08-150-204E-96
12	1517.8	83.9	6039	1	US-08-324-977-11
13	1517.8	83.9	6039	2	US-08-324-977-11
14	1517.8	83.9	6039	3	US-08-324-977-11
15	1517.8	83.9	6039	3	US-08-324-977-11
16	1517.8	83.9	6039	3	US-08-324-977-11
17	1517.8	83.9	6039	3	US-08-324-977-11
18	1517.8	83.9	6039	3	US-08-324-977-11
19	1517.8	83.9	6039	3	US-08-324-977-11
20	1517.8	83.9	6039	3	US-08-324-977-11
21	1517.8	83.9	6039	3	US-08-324-977-11
22	1517.8	83.9	6039	3	US-08-324-977-11
23	1517.8	83.9	6039	3	US-08-324-977-11
24	1517.8	83.9	6039	3	US-08-324-977-11
25	1517.8	83.9	6039	3	US-08-324-977-11
26	1517.8	83.9	6039	3	US-08-324-977-11
27	1517.8	83.9	6039	3	US-08-324-977-11

28	1206.4	66.7	12980	3	US-09-034-756-5	Sequence 5, Appl
29	1204.8	66.6	9646	3	US-08-811-566-1	Sequence 1, Appl
30	1204.8	66.6	9646	3	US-09-034-756-1	Sequence 1, Appl
31	1202	66.4	2082	3	US-08-612-973-47	Sequence 47, Appl
32	1202	66.4	2082	3	US-08-612-973-47	Sequence 47, Appl
33	1201.6	66.4	9599	3	US-09-014-416-2	Sequence 2, Appl
34	1195.2	66.1	9365	4	US-08-432-693-1	Sequence 7, Appl
35	1195.2	66.1	9401	2	US-08-432-693-1	Sequence 1, Appl
36	1195.2	66.1	9416	3	US-08-811-566-19	Sequence 19, Appl
37	1195.2	66.1	9416	3	US-09-034-756-19	Sequence 19, Appl
38	1193.6	66.0	9416	3	US-08-823-895A-26	Sequence 26, Appl
39	1193.6	66.0	9416	4	US-10-104-966-13	Sequence 13, Appl
40	1180.8	65.3	9379	3	US-08-444-818-176	Sequence 176, App
41	1180.8	65.3	9401	5	PCT-US91-02225-9	Sequence 9, Appl
42	1177.6	65.1	8987	3	US-08-444-818-137	Sequence 137, App
43	1177.6	65.1	9185	3	US-08-444-818-122	Sequence 122, App
44	1177.6	65.1	9185	3	US-08-444-818-123	Sequence 123, App
45	1177.6	65.1	9379	3	US-09-388-874-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-08-191-160-21  
; Sequence 21, Application US/08191160  
; Patent No. 6210675  
; GENERAL INFORMATION:  
; APPLICANT: Highfield, Peter Edmund  
; APPLICANT: Rodgers, Brian Colin  
; APPLICANT: Tedder, Richard Seton  
; APPLICANT: Barbara, John Anthony James  
; TITLE OF INVENTION: Viral Agent  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSER: Rothwell, Figg, Ernst & Kurz  
; STREET: 1700 K Street  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage  
; COMPUTER: IBM AT compatible  
; OPERATING SYSTEM: MS-DOS V3.2  
; SOFTWARE: Wordperfect 5.0 (DOS text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/191,160  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/628,516  
; FILING DATE: 17 DEC 1990  
; APPLICATION NUMBER: UK 89 28 562.1  
; FILING DATE: 18 DEC 1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 414.0  
; FILING DATE: 27 FEB 1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 814.1  
; FILING DATE: 03 MAR 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: E. Anthony Figg  
; REGISTRATION NUMBER: 27,195  
; REFERENCE/DOCKET NUMBER: 1645-103A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 833-5740  
; TELEFAX: (202) 833-5740  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2116 base pairs  
; TYPE: nucleotide with corresponding protein  
; STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: cDNA to genomic RNA  
ORIGINAL SOURCE:  
ORGANISM: human; serum infectious for PT-NANBH  
IMMEDIATE SOURCE:  
LIBRARY: cDNA clones from 5' end of the genome  
FEATURES:  
LOCATION: from 308 to 2116 bp start of the PT-NANBH  
LOCATION: polyprotein  
OTHER INFORMATION: viral structural and non-structural  
proteins  
US-08-191-160-21

Query Match 100.0%; Score 1809; DB 3; Length 2116;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCAGATCTTAAACCTCAAGAAAAACCAACGTAACCAACCGCGCCACAG 60  
DB ATGAGCAGATCTTAAACCTCAAGAAAAACCAACGTAACCAACCGCGCCACAG 367  
QY 61 GACGTCAGTCCCGGCGGTGTCAGATCGTTGTGAGTTTACCTGTTCGCGCAGG 120  
DB GACGTCAGTCCCGGCGGTGTCAGATCGTTGTGAGTTTACCTGTTCGCGCAGG 427  
QY 121 GCGCCAGGTTGGTGTGCGCGCACTAGAGAGACTTCGACGCGTCCGAACCTGTGGA 180  
DB GCGCCAGGTTGGTGTGCGCGCACTAGAGAGACTTCGACGCGTCCGAACCTGTGGA 487  
QY 181 AGGCGACAACTATCCCAAGGCTGCGCAGCCGAGGCGAGGCGCTGAGCCCGG 240  
DB AGGCGACAACTATCCCAAGGCTGCGCAGCCGAGGCGAGGCGCTGAGCCCGG 547  
QY 241 TACCTTGGCCCTCTATAGGCAAGAGGCAAGGAGGAGAGTGGCTCTCTGACCC 300  
DB TACCTTGGCCCTCTATAGGCAAGAGGCAAGGAGGAGTGGCTCTCTGACCC 607  
QY 301 CGTGGCTCCCGGCTAGTTGGGCGCCCACTGACCCCGCGCTGAGGTGCGTAAATTGGGT 360  
DB CGTGGCTCCCGGCTAGTTGGGCGCCCACTGACCCCGCGCTGAGGTGCGTAAATTGGGT 667  
QY 361 AAAGTCATGATCCCTCAATGCGGCTTGGCGACCTTCATGAGGATCAATTCGCTCGTC 420  
DB AAAGTCATGATCCCTCAATGCGGCTTGGCGACCTTCATGAGGATCAATTCGCTCGTC 727  
QY 421 GCGCGTCCCTTAGGGGGGCGTGGCCAGGCGCTGGGCGCATGGGTCCGGTTTGGAGAC 480  
DB GCGCGTCCCTTAGGGGGGCGTGGCCAGGCGCTGGGCGCATGGGTCCGGTTTGGAGAC 787  
QY 481 GCGCGTAACTATGCAACAGGGAATTTACCCGGTGTCTCTTCTATCTTCTCTTGGT 540  
DB GCGCGTAACTATGCAACAGGGAATTTACCCGGTGTCTCTTCTATCTTCTCTTGGT 847  
QY 541 TTGCTGTCTGTTTGAACAATTCACGCTTCCGTTATGAAGTCGCAACGTCGCCGATC 600  
DB TTGCTGTCTGTTTGAACAATTCACGCTTCCGTTATGAAGTCGCAACGTCGCCGATC 907  
QY 601 TACCAATGCAAGAAAGATTGCTCCAACTCAAGCATTCGTGTAGAGAACGCGGACATGATC 660  
DB TACCAATGCAAGAAAGATTGCTCCAACTCAAGCATTCGTGTAGAGAACGCGGACATGATC 967  
QY 661 ATGCAACCCCGGGGTGTGTGCTGTGTGCGGGAGGGTAAATTCCTCCGCTGCTGGGTA 720  
DB ATGCAACCCCGGGGTGTGTGCTGTGTGCGGGAGGGTAAATTCCTCCGCTGCTGGGTA 1027  
QY 721 GCGCTCACTCCACGCTCGCGGCAAGAGCGCAGCATTCGCCACTGCGCAATACGAGCG 780  
DB GCGCTCACTCCACGCTCGCGGCAAGAGCGCAGCATTCGCCACTGCGCAATACGAGCG 1087  
QY 781 CACGTCGATTTGCTGTGGGGCGGCTGCTTCTGCTCGCATATGATGTGGGGATCTC 840  
DB CACGTCGATTTGCTGTGGGGCGGCTGCTTCTGCTCGCATATGATGTGGGGATCTC 1147

QY 841 TCGCGATCTGTTTCTCTGCTCTCAGCTGTCACTTCTCGCTCGCGCATCAGACG 900  
DB TCGCGATCTGTTTCTCTGCTCTCAGCTGTCACTTCTCGCTCGCGCATCAGACG 1207  
QY 901 GTACAGAGCTGCAATTGTTCAATCTATCCGCGCAAGTATCAGGTCAACGCAATGGCTTGG 960  
DB GTACAGAGCTGCAATTGTTCAATCTATCCGCGCAAGTATCAGGTCAACGCAATGGCTTGG 1267  
QY 961 GATATGAGATGAATCTGTGCACTTACAGCAGCCCTAGTGGTATGCGAGCTACTCCGATC 1020  
DB GATATGAGATGAATCTGTGCACTTACAGCAGCCCTAGTGGTATGCGAGCTACTCCGATC 1327  
QY 1021 CCACAAGCTGTCTGTGACATGTGTGCGGCGGCCCACTGCGGAGTCTTGGCGGCTTGGC 1080  
DB CCACAAGCTGTCTGTGACATGTGTGCGGCGGCCCACTGCGGAGTCTTGGCGGCTTGGC 1387  
QY 1081 TACTATTCATGTGTGGGAACTGGGCTTAAGTCTTGTGTGATGCTACTCTTGGCGGC 1140  
DB TACTATTCATGTGTGGGAACTGGGCTTAAGTCTTGTGTGATGCTACTCTTGGCGGC 1447  
QY 1141 GTTGAAGGGGAACTTACAGACAGGGGGGACACACGCGCGCGCGCCACGCGGCTTACA 1200  
DB GTTGAAGGGGAACTTACAGACAGGGGGGACACACGCGCGCGCGCCACGCGGCTTACA 1507  
QY 1201 TCCCTCTTCAACCTTGGCGCGCTCAGAAAAATCCAGCTTGTAAACACCAACGCGACGTGG 1260  
DB TCCCTCTTCAACCTTGGCGCGCTCAGAAAAATCCAGCTTGTAAACACCAACGCGACGTGG 1567  
QY 1261 CACATCAAAGAACTGCTTGAATCTGCAATGATCTCTCCAACTGGGTTCTTGGCGCG 1320  
DB CACATCAAAGAACTGCTTGAATCTGCAATGATCTCTCCAACTGGGTTCTTGGCGCG 1627  
QY 1321 CTGTTCTAACAGCAAGGTTCAATGCGGTGCGGAGTCTCAGAGGCGCATAGGCGACGCGC 1380  
DB CTGTTCTAACAGCAAGGTTCAATGCGGTGCGGAGTCTCAGAGGCGCATAGGCGACGCGC 1687  
QY 1381 CCCATTGACCAATTGCAATCAGAGGGTGGGTCCTCATCTTAATTAATGAGTCCACGCGCTTG 1440  
DB CCCATTGACCAATTGCAATCAGAGGGTGGGTCCTCATCTTAATTAATGAGTCCACGCGCTTG 1747  
QY 1441 GACCAAGGCGCTTATGCTGGCACTAGGCACTTCAACCGTGTGATCTGTCGCGCTTG 1500  
DB GACCAAGGCGCTTATGCTGGCACTAGGCACTTCAACCGTGTGATCTGTCGCGCTTG 1807  
QY 1501 CAGGTGTGCGCCAGGTGATCTGTTTCACTCAAGCCCTGTTGTGTGGGAGACACGAT 1560  
DB CAGGTGTGCGCCAGGTGATCTGTTTCACTCAAGCCCTGTTGTGTGGGAGACACGAT 1867  
QY 1561 CGTTTCGCGCGCCCTTACGTAAGATGGGATGAGATGAGACGATGCTCTTCAAC 1620  
DB CGTTTCGCGCGCCCTTACGTAAGATGGGATGAGATGAGACGATGCTCTTCAAC 1927  
QY 1621 AACACGCGGCGCGCAACGCGGCACTGGTTCGGCTGTACATGATGAATAGCAACGCGGCTTC 1680  
DB AACACGCGGCGCGCAACGCGGCACTGGTTCGGCTGTACATGATGAATAGCAACGCGGCTTC 1987  
QY 1681 ACCAAGACGTGTGGGGGCGCCCGGTGCAACATCGGGGGGGTGGCAACAACCTTTGATC 1740  
DB ACCAAGACGTGTGGGGGCGCCCGGTGCAACATCGGGGGGGTGGCAACAACCTTTGATC 2047  
QY 1741 TGCCCAACGAGCTCTTCCGAGAACATCCGAGGCACTTACACCAATATGCGGCTTGGGG 1800  
DB TGCCCAACGAGCTCTTCCGAGAACATCCGAGGCACTTACACCAATATGCGGCTTGGGG 2107

RESULT 2  
US-09-539-601-1  
; Sequence 1, Application US/09539601C  
; Patent No. 6630343

```
GENERAL INFORMATION:
APPLICANT: Bartenschlager, Ralf FW
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
FILE REFERENCE: all sequences
CURRENT APPLICATION NUMBER: US/09/539,601C
EARLIER FILING DATE: 2001-08-30
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 11076
TYPE: DNA
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)..(341)
OTHER INFORMATION: construct 1389/Core-3'/wt
FEATURE:
NAME/KEY: CD3
LOCATION: (342)..(1193)
OTHER INFORMATION: hepatitis C virus core - neomycin
OTHER INFORMATION: phosphotransferase fusion protein
FEATURE:
NAME/KEY: RBS
LOCATION: (1302)..(1812)
OTHER INFORMATION: internal ribosome entry site from
OTHER INFORMATION: encephalomyocarditis virus
FEATURE:
NAME/KEY: CD3
LOCATION: (1813)..(10845)
OTHER INFORMATION: hepatitis C virus polypeptide from core to
OTHER INFORMATION: nonstructural protein NS5B; parental sequence
OTHER INFORMATION: without cell culture-adaptive mutations
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (10846)..(11076)
US-09-539-601-1

Query Match      86.8%; Score 1570.6; DB 4; Length 11076;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY      1  ATGAGCAGCAATCTCTAACTCTCAAGAAACCAACGTTACACCAACGCGCGCCACAG 60
DB      1813  ATGAGCAGCAATCTCTAACTCTCAAGAAACCAACGTTACACCAACGCGCGCCACAG 1872
QY      61  GACGTCAAGTCCCGGCGGCGGTGTCAGATCGTTGGTGGATTACCTGTTGCCGCGCAG 120
DB      1873  GACGTCAAGTCCCGGCGGCGGTGTCAGATCGTTGGTGGATTACCTGTTGCCGCGCAG 1932
QY      121  GGGCCGAGGTTGGTGTGCGCGCATGAGAAAGACTTCCGAGCGGTCCGAACTCTGTGA 180
DB      1933  GGGCCGAGGTTGGTGTGCGCGCATGAGAAAGACTTCCGAGCGGTCCGAACTCTGTGA 1992
QY      193  GGGCCGAGGTTGGTGTGCGCGCATGAGAAAGACTTCCGAGCGGTCCGAACTCTGTGA 1992
DB      1933  GGGCCGAGGTTGGTGTGCGCGCATGAGAAAGACTTCCGAGCGGTCCGAACTCTGTGA 1992
QY      241  TACCCCTGGGCGCTCTATGAGCAACGAGGCGATGGGGGAGAGATGGGCTCGTCAACC 300
DB      1993  TACCCCTGGGCGCTCTATGAGCAACGAGGCGATGGGGGAGAGATGGGCTCGTCAACC 2052
QY      241  TACCCCTGGGCGCTCTATGAGCAACGAGGCGATGGGGGAGAGATGGGCTCGTCAACC 300
DB      2053  TACCCCTGGGCGCTCTATGAGCAACGAGGCGATGGGGGAGAGATGGGCTCGTCAACC 2112
QY      301  CGTGGCTCCCGGCTTATGAGGCGCCCACTGACCCCGGCGTATGATCGCGTAATTTGG 360
DB      2113  CGTGGCTCCCGGCTTATGAGGCGCCCACTGACCCCGGCGTATGATCGCGTAATTTGG 2172
QY      361  AAGTATGATGATCTCTCATGATGCGGCTTCCGCGACCTCATGGGGTATCATTTCCG 420
DB      2173  AAGTATGATGATCTCTCATGATGCGGCTTCCGCGACCTCATGGGGTATCATTTCCG 2232
QY      421  GGGCTCCCTTATGAGGCGGCGTCCAGAGGCGCTGGCGCATGGCGTCCGGGTTTGA 480
DB      421  GGGCTCCCTTATGAGGCGGCGTCCAGAGGCGCTGGCGCATGGCGTCCGGGTTTGA 480
```

```
DB      2233  GGGCGCCCTCTATGAGGCGGCGTCCAGGCGCTTGGCGCATGGCGTCCGGTTCTGAGAGAC 2292
QY      481  GGGCGTGAATATGCAACAGGGAATTTACCCGGTGTCTTTCTCTATCTTCTCTTGGGT 540
DB      2293  GGGCGTGAATATGCAACAGGGAATCTGCGCGGTCTCTCTTTCTCTATCTTCTTGGGT 2352
QY      541  TTGCTGTCTGTTTGGCAATTCAGCTTCGCTTATGAGTGCAGCAAGTGTCCGGATC 600
DB      2353  TTGCTGTCTGTTTGGCAATTCAGCTTCGCTTATGAGTGCAGCAAGTGTCCGGATC 2412
QY      601  TACCATGTCCAGCAACGATTTCTCAATCAAGCATGTGTACAGACAGCGGACATGATC 660
DB      2413  TACCATGTCCAGCAACGATTTCTCAATCAAGCATGTGTACAGACAGCGGACATGATC 2472
QY      661  ATGCAACACCCCGGCGTGTGTGCTGTGTCGGGAGGTTATTTCTTCCGCTGTGGTGA 720
DB      2473  ATGCAATACCCCGGCGTGTGTGCTGTGTCGGGAGAACATCTCTCCGCTGTGGGTA 2532
QY      721  GGGCTCACTCCGACGCTGCGGCGCAAGAGCGCAGATCCCACTGCGCAATACGACG 780
DB      2533  GGGCTCACTCCGACGCTGCGGCGCAAGAGCGCAGATCCCACTGCGCAATACGACG 2592
QY      781  CAGTGTGATTTGCTGTGGGCGGCTGTCTGTCTCCGCTATGATGATGTGGGAGATCTC 840
DB      2593  CAGTGTGATTTGCTGTGGGCGGCTGTCTGTCTCCGCTATGATGATGTGGGAGATCTC 2652
QY      841  TGGGATCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
DB      2653  TGGGATCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2712
QY      901  GTTACAGGATGCAATTTGTTCAATCTATCCCGGCGCAGTATCAGATCACCGATGCTTG 960
DB      2713  GTTACAGGATGCAATTTGTTCAATCTATCCCGGCGCAGTATCAGATCACCGATGCTTG 2772
QY      961  GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB      2773  GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2832
QY      1021  CCAACAAGCTGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB      2833  CCAACAAGCTGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2892
QY      1081  TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB      2893  TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2952
QY      1141  GTTGAACGAGGAACTTACAGACAGAGGAGGACACAGCGCGCGCCACCGGCTTACA 1200
DB      2953  GTTGAACGAGGAACTTACAGACAGAGGAGGACAGAGGAGGACAGAGGAGGATTAAG 3012
QY      1201  TCCCTCTTCAACCTTGGGCGGCTCAAGAAATTCAGCTTGTAAACACCAAGCGAGCTGG 1260
DB      3013  TCCCTCTTCAACCTTGGGCGGCTCAAGAAATTCAGCTTGTAAACACCAAGCGAGCTGG 3072
QY      1261  CACATCAACAGAACTGCTTGAATGCAATGATCTCCCTCAAACTGGGTTCTTGGCGG 1320
DB      3073  CACATCAACAGAACTGCTTGAATGCAATGATCTCCCTCAAACTGGGTTCTTGGCGG 3132
QY      1321  CTGTTCTACACGACAGGTTCAATGCTGCTCGATGCTCAGAGCGATGCGACGCTGCG 1380
DB      3133  CTGTTCTACACGACAGGTTCAATGCTGCTCGATGCTCAGAGCGATGCGACGCTGCG 3192
QY      1381  CCCATTGACCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
DB      3193  CCCATTGACCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3252
QY      1441  GACGAGGCGCTTATGCTGAGCACTACGACCTCAACGATGATGATGATGATGATGATG 1500
DB      3253  GACGAGGCGCTTATGCTGAGCACTACGACCTCAACGATGATGATGATGATGATGATG 3312
QY      1501  CAGGTGTGTGACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB      3313  CAGGTGTGTGACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3372
```



```
|||||
Db 2953 GTTGAAGGGGGAACCTATGACAGAGGGGAGACATGCGCAAAAACACCTCGGGGATTACG 3012
Qy 1201 TCCCTCTTCAACCTGAGGCGGGCTCAAAAATCAGCTGTGTAACCAACAGGAGCTGG 1260
Db 3013 TCCCTCTTCAACCGGGGCTATCCAGAAATCAGCTGTGTAACCAACAGGAGCTGG 3072
Qy 1261 CACATCAAGAACTGGCCTTGAACCTGCAATGATCTCCCTCAACCTGGTCTTGGCGG 1320
Db 3073 CACATCAAGAACTGGCCTTGAACCTGCAATGATCTCCCTCAACCTGGTCTTGGCGG 3132
Qy 1321 CTGTTCTACAGCAGAGTTCAATGAGTCGAGATGCTCAGAGGCGCATGGCAGCTGCCG 1380
Db 3133 CTGTTCTACAGCAGAGTTCAATGAGTCGAGATGCTCAGAGGCGCATGGCAGCTGCCG 3192
Qy 1381 CCCATTGACCAAGTTGATGACAGGGGAGGGGCTCCATCATTAATGATGCTCCAGGCTTG 1440
Db 3193 CCCATTGACCAAGTTGATGACAGGGGAGGGGCTCCATCATTAATGATGCTCCAGGCTTG 3252
Qy 1441 GACGAGAGGCGCTTATGCTGCTGCACTACGCACTCCGATGCTGCTGCTGCTGCTGCTG 1500
Db 3253 GACGAGAGGCGCTTATGCTGCTGCACTACGCACTCCGATGCTGCTGCTGCTGCTGCTG 3312
Qy 1501 CAGGTGTGAGCCCAAGTGTACTGTTCACTCAAGCCCTGTTGTTGTTGTTGTTGTTGTTG 1560
Db 3313 CAGGTGTGAGCCCAAGTGTACTGTTCACTCAAGCCCTGTTGTTGTTGTTGTTGTTGTTG 3372
Qy 1561 CGTTTCGAGCGCCCTAAGTACAGATGAGGAGTGAAGTGAAGCGAGCTGCTGCTTCAAC 1620
Db 3373 CGTTTCGAGCGCCCTAAGTACAGATGAGGAGTGAAGTGAAGCGAGCTGCTGCTTCAAC 3432
Qy 1621 AACACGCGGCGCCGCAAGGAGCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db 3433 AACACGCGGCGCCGCAAGGAGCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3492
Qy 1681 ACCAAGACGTTGAGGAGGCGCCCGCTGCAACATCGAGGAGGCTCGGCAACCACTTTGATC 1740
Db 3493 ACCAAGACGTTGAGGAGGCGCCCGCTGCAACATCGAGGAGGCTCGGCAACCACTTTGATC 3552
Qy 1741 TGCCCAAGCACTGCTTCCGGAAGCATCCGAGGCACTTACCAACCAATCGGTTGCGG 1800
Db 3553 TGCCCAAGCACTGCTTCCGGAAGCATCCGAGGCACTTACCAACCAATCGGTTGCGG 3612
Qy 1801 CCTTGGTTG 1809
Db 3613 CCTTGGTTG 3621

RESULT 4
US-09-539-601-25
; Sequence 25, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Barteneschlager, Ralf FM
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/5.1
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
```

```
OTHER INFORMATION: hepatitis C virus core - neomycin
OTHER INFORMATION: phosphotransferase fusion protein
;
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polypeptide from core to
; OTHER INFORMATION: nonstructural protein NS5B, carries cell
; OTHER INFORMATION: culture-adaptive mutations of clone 5.1
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
US-09-539-601-25

Query Match 86.8%; Score 1570.6; DB 4; Length 11076;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

Qy 1 ATGAGCAGCAATCTTAAACCTCAAGAAACCAACGTAACCAACCGCGCCGACAG 60
Db 1813 ATGAGCAGCAATCTTAAACCTCAAGAAACCAACGTAACCAACCGCGCCGACAG 1872
Qy 61 GACGTAAAGTTCCCGGCGGCTGTCAGATCGTTGAGAGTTTACTTGTTCGCGCAGG 120
Db 1873 GACGTAAAGTTCCCGGCGGCTGTCAGATCGTTGAGAGTTTACTTGTTCGCGCAGG 1932
Qy 121 GGGCCAGGTTGGGAGTGGCGGCACTAGAAAGACTTCCGAGCGTGCACCTCGTGA 180
Db 1933 GGGCCAGGTTGGGAGTGGCGGCACTAGAAAGACTTCCGAGCGTGCACCTCGTGA 1992
Qy 181 AGCGCAACCTATCCCAAGCTCGCAAGCCGAGGAGGAGGCTGAGCTCAGCCGAG 240
Db 1993 AGCGCAACCTATCCCAAGCTCGCAAGCCGAGGAGGAGGCTGAGCTCAGCCGAG 2052
Qy 241 TACCTTGGCCCTCTATAGGCAACGAGGAGTGGGAGGAGTGGGAGTGGGAGTGGGAG 300
Db 2053 TACCTTGGCCCTCTATAGGCAACGAGGAGTGGGAGGAGTGGGAGTGGGAGTGGGAG 2112
Qy 301 CGTGGCTCCGCGCTGTTGAGGAGCCCACTGACCCCGGAGTGGGAGTGGGAGTGGGAG 360
Db 2113 CGTGGCTCCGCGCTGTTGAGGAGCCCACTGACCCCGGAGTGGGAGTGGGAGTGGGAG 2172
Qy 361 AAAGTCATGATACCTCTCATGCGGCTTGGCGAGCTCATGAGGATGATTCGCTGCT 420
Db 2173 AAAGTCATGATACCTCTCATGCGGCTTGGCGAGCTCATGAGGATGATTCGCTGCT 2232
Qy 421 GGGCTCCCTTAGGGGAGGCTGCGAGGAGCTTGGGAGGAGTGGGAGTGGGAGTGGGAG 480
Db 2233 GGGCTCCCTTAGGGGAGGCTGCGAGGAGCTTGGGAGGAGTGGGAGTGGGAGTGGGAG 2292
Qy 481 GGGGTAACATGCAACAGGAAATTACCGGTTGCTTCTCTATCTTCTCTTGGCT 540
Db 2293 GGGGTAACATGCAACAGGAAATTACCGGTTGCTTCTCTATCTTCTCTTGGCT 2352
Qy 541 TTGCTGCTGTTTGAACATTCAGCTTCCGTTATGAGTGGGAGTGGGAGTGGGAGTGG 600
Db 2353 TTGCTGCTGTTTGAACATTCAGCTTCCGTTATGAGTGGGAGTGGGAGTGGGAGTGG 2412
Qy 601 TACATGTCAAGCAAGATTTGCTCAACTCAAGCATGTGTAGAGACAGCGACATGATC 660
Db 2413 TACATGTCAAGCAAGATTTGCTCAACTCAAGCATGTGTAGAGACAGCGACATGATC 2472
Qy 661 ATGCAACCGCGGAGTGGGAGCTGTCGCGGAGGAGTAAATCTCCGCTGCTGGGTA 720
Db 2473 ATGCAACCGCGGAGTGGGAGCTGTCGCGGAGGAGTAAATCTCCGCTGCTGGGTA 2532
Qy 721 GGGCTCACTCCACGCTGCGGCAAGAGCGCGAGATCCCACTGCGCAATATGACGCG 780
Db 2533 GGGCTCACTCCACGCTGCGGCAAGAGCGCTAGGATCCCACTGCGCAATATGACGCG 2592
```







[illegible]

```

QY 1381 CCCATTGACCAAGTTTCGATCAGGGGTGGGGTCCCATCACTATATATGATGCCAGGCTTG 1440
Db 1722 CCCATTGACTGTGGTTCGCCAAGGGGTGGGGCCCATCACTATATGAGGCCCTTAACAGCTCG 1781
QY 1441 GACCAGAGAGCCCTTATTTGCTGGCACTAGCAACCTCAACCGTGTGTATCGTCCCGCGTTC 1500
Db 1782 GATCAGAGGCTTATTTGCTGGCACTTAGCGCGCTGCACCGTGTGTGTGTATCGTCCCGCGTTC 1841
QY 1501 CAGGTGTGTGGCCCAAGTGTACTGTTTCACTTCAAGCCCTGTTGTGTGGGAGCAACGAT 1560
Db 1842 CAGGTGTGTGTCCAGTGTATTTGTTTCAACCCCAAGCCCTGTTGTGTGGGAGCAACGAT 1901
QY 1561 CGTTTCGGCGCCCTTACGTAAGATGATGGGTGAGAAATGAGACGAGCGTGTCTTCTCAAC 1620
Db 1902 CGTTTCGGGTGTCCCTTACGTAAGATGATGGGTGAGAAATGAGACGAGTGTCTTCTCAAC 1961
QY 1621 AACACGCGCGCCGCAACGAGGCAACTGTGTCGGCTGTACATGATGAAATAGCACCGAGGTTTC 1680
Db 1962 AACACGCGCTCCGCAACGAGGCAACTGTGTCGGCTGTACATGATGAAATAGCACTCGGGTTC 2021
QY 1681 ACCAAGACGTGTGGGGGCCCCCGTGCAACATCGGGGGGGTTCGCAACACACACTTTGATC 1740
Db 2022 ACTAAGACGTGTGGAGGTCGCCCGCTGTATACATCGGGGGGGTTCGTTAACCGCACCTTATTC 2081
QY 1741 TGCCCCACGCACTGTTCCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTTCGGGG 1800
Db 2082 TGCCCCACGCACTGTTCCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTTCGGGG 2141
QY 1801 CCTTGCTTG 1809
Db 2142 CCTTGCTTG 2150

RESULT 7
US-09-014-416-6
; Sequence 6, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masaaki
; APPLICANT: Buhn, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; CURRENT FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 9599
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-014-416-6

Query Match 86.0%; Score 1556.2; DB 3; Length 9599;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY 1 ATGAGCAGAGATCTCTTAACTCAAGAAAAACAACGTAACACCAACCGCGCCACAG 60
Db 342 ATGAGCAGAGATCTCTTAACTCAAGAAAAACAACGTAACACCAACCGCGCCACAG 401
QY 61 GACGTCAAGTTCCCGGGCGGTGTGATCGTGTGTGAGTTTACTGTTCGCGCAGG 120
Db 402 GACGTCAAGTTCCCGGGCGGTGTGATCGTGTGTGAGTTTACTGTTCGCGCAGG 461
QY 121 GGGCCAGAGTTGGGTGTGCGCGGCACCTAGCAACTTCCGAGCGGTGCCAAGCTGTGGA 180
Db 462 GGGCCAGAGTTGGGTGTGCGCGGCACCTAGCAAGGCTTCCGAGCGGTGCCAAGCTGTGGA 521

```

181 AGGCGACAACTATCCCAAGGCTGGCGAGCCCGGAGGCGCTGGGCTGAGCCCGG 240  
522 AGCGGACAACTATCCCAAGGCTGGCGAGCCCGGAGGCGCTGGGCTGAGCCCGG 581  
241 TACCCCTGGGCGCTCTATGAGCAAGAGGCGATGGGGTGGGAGAGATGCTCTGTACCC 300  
582 TACCCCTGGGCGCTCTATGAGCAAGAGGCGCTGGGGTGGGAGAGATGCTCTGTACCC 641  
301 CGTGGCTCCCGGCTATGTTGGGGCGCCCACTGACCCCGGCGTAAAGTGGTATTTGGGT 360  
642 CGCGGCTCCCGGCTATGTTGGGGCGCCCACTGACCCCGGCGTAAAGTGGTATTTGGGT 701  
361 AAGATATGATATCCCTCAATGCGGCTTCGCGGACTGATGGGGTCAATTCGCGCTGTC 420  
702 AAGATATGATATCCCTCAATGCGGCTTCGCGGACTGATGGGGTCAATTCGCGCTGTC 761  
421 GCGGCTCCCTTAGGGGGCGCTGCGAGGCGCTGGCGCATGCGCTCGGGTCTGAGAGAC 480  
762 GCGGCGCCCTTAGGGGGCGCTGCGAGGCGCTGGCGCATGCGCTCGGGTCTGAGAGAC 821  
481 GCGGCTATGATGCAACAGGGAATTTAACCGGCTGCTCTTTCTCTATCTTCTTGCT 540  
822 GCGGCTATGATGCAACAGGGAATTTAACCGGCTGCTCTTTCTCTATCTTCTTGCT 881  
541 TGGCTGCTCTGTTGACCAATTCAGCTTCGCTTATGAAAGTGGGCAAGCTTCGGGATC 600  
882 CTGCTGCTCTGTTGACCAATTCAGCTTCGCTTATGAAAGTGGGCAAGCTTCGGGATC 941  
601 TACCATGTCAGCAAGATGCTCTCAACTCAAGCATGCTGTACAGACAGCGGATGATC 660  
942 TACCATGTCAGCAAGATGCTCTCAACTCAAGCATGCTGTGTATGAGGCAAGCGGATGATC 1001  
661 ATGCAACACCCCGGGGTGGTGGCTGTGTGTCGGGAGGGTAAATCTTCGCGCTGGGATC 720  
1002 ATGCAACACCCCGGGGTGGTGGCTGTGTGTCGGGAGGGTAAACAGCTCCGCTGGGATC 1061  
721 GGGCTACCTCCAGCGCTCGCGGCGAGGAGCGGAGATCCCACTGCGCAATTAAGACGC 780  
1062 GGGCTACCTCCAGCGCTCGCGGCGAGGAGATCCCACTGCGCAATTAAGACGC 1121  
781 CAGCTGATTTGCTGCTGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
1122 CAGCTGATTTGCTGCTGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181  
841 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
1182 TGGGATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241  
901 GTACAGGATGCAATTTGCTCAATCTATCCCGGCGAGTACAGGTACCGCATGGCTGG 960  
1242 GTACAGGATGCAATTTGCTCAATCTATCCCGGCGAGTACAGGTACCGCATGGCTGG 1301  
961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
1302 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361  
1021 CCACAAAGCTGTGTGAGCAATGTGTGCGGGGCGCCACTGGGAGTCTTGGGCGGCTTGGC 1080  
1362 CCACAAAGCTGTGTGAGCAATGTGTGCGGGGCGCCACTGGGAGTCTTGGGCGGCTTGGC 1421  
1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
1422 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1481  
1141 GTTGAAGGGAACCTTACAGCAAGAGGGAACACAGCGCGCGCGCCCAAGGCTTAC 1200  
1482 GTTGAAGGGAACCTTACAGCAAGAGGGAACAGCGCGCGCGCCCAAGGCTTAC 1541  
1201 TCCCTCTTCAACACTGGGCGGCTCAAGAAATCAGCTTGTAAACCAACGAGCACTGG 1260  
1542 TCCCTCTTCAACACTGGGCGGCTCAAGAAATCAGCTTGTGTAAACCAACGAGCACTGG 1601

1261 CACATCAACAGAACTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1602 CACATCAACAGAACTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1661  
1321 CTGTTCTACAGCAAGGTTCAATGCTGCGATGCTGAGCGGATGCGGATGCGGATGCGG 1380  
1662 CTGTTCTACAGCAAGGTTCAATGCTGCGATGCTGAGCGGATGCGGATGCGGATGCGG 1721  
1381 CCCATTTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
1722 CCCATTTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1781  
1441 GACCAAGGCGCTTATTTGCTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAG 1500  
1782 GATCAGAGGCTTATTTGCTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAG 1841  
1501 CAGGTGTGTGCTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAG 1560  
1842 CAGGTGTGTGCTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAG 1901  
1561 CGTTTGGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
1902 CGTTTGGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1961  
1621 AACAGCGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
1962 AACAGCGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2021  
1681 ACCAAGCGTGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1740  
2022 ACTAAGCGTGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2081  
1741 TGGCCCAAGGATGCTTCCGGAAGCATCCCGAGGCGCACTTAAACCAAAATGGCTTGG 1800  
2082 TGGCCCAAGGATGCTTCCGGAAGCATCCCGAGGCGCACTTAAACCAAAATGGCTTGG 2141  
1801 CCTGGTTG 1809  
2142 CCTGGTTG 2150

RESULT 8  
US-09-827-688-6  
; Sequence 6, Application US/09827688  
; Patent No. 6821955  
; GENERAL INFORMATION:  
; APPLICANT: ORSON, FRANK  
; APPLICANT: KINSEY, BERNIA  
; APPLICANT: BHOGLI, BALBIR  
; TITLE OF INVENTION: MACROAGGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION  
; FILE REFERENCE: P01949US1/10004014  
; CURRENT APPLICATION NUMBER: US/09/827,688  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: 60/195,680  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 9413  
; TYPE: DNA  
; ORGANISM: HEPATITIS C  
US-09-827-688-6

Query Match 85.3%; Score 1543.4; DB 4; Length 9413;  
Best Local Similarity 90.8%; Pred. No. 0;  
Matches 1643; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

1 ATGAGCAGCAATCTTAACTCAAAAGAAACCAAAAGTAAACCAACGAGCGCGCCACAG 60  
330 ATGAGCAGCAATCTTAACTCAAAAGAAACCAAAAGTAAACCAACGAGCGCGCCACAG 389  
61 GAGGTCAAGTCCCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 120

Db 390 GACGTTAAGTCCCGGCGGTGTGAGATCGTTGGAGATTACCTGTTGCCGCGCAGG 449  
Qy 121 GGGCCCAAGTTGGGTGTGCGCGCACTAGAGAGACTCCGACGGGTCCGAACCTGTGGA 180  
Db 450 GGGCCCAAGTTGGGTGTGCGCGCACTAGAGAGACTCCGACGGGTCCGAACCTGTGGA 509  
Qy 181 AGGCGCAACACTATCCCGAAGGCTGCGCAGCGGCGAGGGGCTGAGCCCGG 240  
Db 510 AGGCGCAACACTATCCCGAAGGCTGCGCAGCGGCGAGGGGCTGAGCCCGG 569  
Qy 241 TACCTTGGGCTCTCTATGCGCAACGAGGCGATGAGGTGGGAGAGATGCTCTGACCC 300  
Db 570 TACCTTGGGCTCTCTATGCGCAACGAGGCGATGAGGTGGGAGAGATGCTCTGACCC 629  
Qy 301 CGTGCTCCCGGCTAGTTGGGGGCGCGACCTGACCCCGCGTAAAGTTCGCTATTTGGGT 360  
Db 630 CGTGCTCCCGGCTAGTTGGGGGCGCGACCTGACCCCGCGTAAAGTTCGCTATTTGGGT 689  
Qy 361 AAAGTCATGATACCTCATGATGCGGCTTGCAGCACTCATGAGGGTACATTCGCTGTC 420  
Db 690 AAAGTCATGATACCTCATGATGCGGCTTGCAGCACTCATGAGGGTACATTCGCTGTC 749  
Qy 421 GGGCTTCCCTTAGGGGGGCGCTGCCAGGGCTTGGGCGATGCGTCCGGTCTGAGAGAC 480  
Db 750 GGGCTTCCCTTAGGGGGGCGCTGCCAGGGCTTGGGCGATGCGTCCGGTCTGAGAGAC 809  
Qy 481 GGGCGTAATGACAAAGGGGAATTTACCCGGTGTCTCTTCTCATCTTCTCTGCT 540  
Db 810 GGGCGTAATGACAAAGGGGAATTTACCCGGTGTCTCTTCTCATCTTCTCTGCT 869  
Qy 541 TTGCTGTCTGTTTACCACTTCAGCTTCGCTTATGAGTCCGCAAGCTGTCGGGATC 600  
Db 870 TTGCTGTCTGTTTACCACTTCAGCTTCGCTTATGAGTCCGCAAGCTGTCGGGATC 929  
Qy 601 TACCACTGTCAGCAAGATTTGCTTCAACTAGCACTGTGTACGAGACAGCGGACATGATC 660  
Db 930 TACCACTGTCAGCAAGATTTGCTTCAACTAGCACTGTGTACGAGACAGCGGACATGATC 989  
Qy 661 ATGCAACACCCCGGGGTGTGCTGCTGTCGGGAGGGGAATTTCTCCCGCTGCTGGGTA 720  
Db 990 ATGCAACACCCCGGGGTGTGCTGCTGTCGGGAGGGGAATTTCTCCCGCTGCTGGGTA 1049  
Qy 721 GGGCTCACTCCCGCAGCTGCGCGCAAGAGCGCCAGCATCCCACTGCGCAATACAGAGC 780  
Db 1050 GGGCTCACTCCCGCAGCTGCGCGCAAGAGCGCCAGCATCCCACTGCGCAATACAGAGC 1109  
Qy 781 CACGTGATTTGCTGTTGGGGGCGCTGCTTCTGCTCCGCTATGATGATGGGGATCTC 840  
Db 1110 CACGTGATTTGCTGTTGGGGGCGCTGCTTCTGCTCCGCTATGATGATGGGGATCTC 1169  
Qy 841 TGGCGATGTTTCTGCTGCTCCTGAGCGCTTCACTTCTGCGCTGCGCGACATACAGC 900  
Db 1170 TGGCGATGTTTCTGCTGCTCCTGAGCGCTTCACTTCTGCGCTGCGCGACATACAGC 1229  
Qy 901 GTACAGAGCTGCAATTTGCTCAATCTATCCCGCGCAAGTATCAGGTACCGCATGCTTGG 960  
Db 1230 GTACAGAGCTGCAATTTGCTCAATCTATCCCGCGCAAGTATCAGGTACCGCATGCTTGG 1289  
Qy 961 GATATGATGATGAATCTGCTCACTTACAGAGCGCTTATGATGATGAGCTATCCGATC 1020  
Db 1290 GATATGATGATGAATCTGCTCACTTACAGAGCGCTTATGATGATGAGCTATCCGATC 1349  
Qy 1021 CCACAAGCTGTGTGAGACATGTTGGGCGCGCGCACTGGGAGGTCTGGGCGGCTTGGC 1080  
Db 1350 CCACAAGCTGTGTGAGACATGTTGGGCGCGCGCACTGGGAGGTCTGGGCGGCTTGGC 1409  
Qy 1081 TACTATTCCATGTTGGGGAAGCTGGGCTTAAAGTCTTGTGTGATGATCTCTTTGGCCG 1140  
Db 1410 TACTATTCCATGTTGGGGAAGCTGGGCTTAAAGTCTTGTGTGATGATCTCTTTGGCCG 1469  
Qy 1141 GTTGAAGGGGAAGCTTTACAGACAGGGGGGACACAGCGCGCGCCCAAGGCTTTACA 1200

Db 1470 GTTGAAGGGGACACCCCACTGATACAGGGGGAAGGGTAGCTTCACAGACCAGAGCTCTG 1529  
Qy 1201 TCCCTTTACACCTGGGCGGGCTCAGAAATCCAGCTTGTAAACACCAAGGGAGCTGG 1260  
Db 1350 TCTGGCTTCAAGAGGCCATCTCAGAAATCCAACTGTTGAACACCAAGGGAGCTGG 1589  
Qy 1261 CACATCAACAGAGCTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
Db 1590 CACATCAACAGAGCTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1649  
Qy 1321 CTGTTCTACAGCAAGGTTCAATGCTCCGAGTCTCAGAGCGGATGAGCGGATGAGCGG 1380  
Db 1650 CTGTTCTACAGCAAGGTTCAATGCTCCGAGTCTCAGAGCGGATGAGCGGATGAGCGG 1709  
Qy 1381 CCATTTGACAGTTGATCAGAGGGGAGGGTCCCATCACTTATTAATGATGATGATGATGAT 1440  
Db 1710 CCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1769  
Qy 1441 GACCAAGGCGCTTATTTGCTGACCTACAGCACTCAACCGTGTGATGATGATGATGATGAT 1500  
Db 1770 GACCAAGGCGCTTATTTGCTGACCTACAGCACTCAACCGTGTGATGATGATGATGATGAT 1829  
Qy 1501 CAGGTGTGAGCCCAAGTGTATCTGTTCACTCAAGCCCTGTTGTGTGGGAGCAAGCAT 1560  
Db 1830 CAGGTGTGAGCCCAAGTGTATGCTTCACTCAAGCCCTGTTGTGTGGGAGCAAGCAT 1889  
Qy 1561 GCTTTCGCGCGCCCTTACAGTACAGATGAGGGGTGAGATGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1890 GCTTTCGCGCGCCCTTACAGTGTAGCTGGGGGAGATGAGAGAGAGAGAGAGAGAGAGAG 1949  
Qy 1621 AACACGCGCGCCCAAGGCGCAAGTGTGCTGCTGTATCATGATGATGATGATGATGATGATG 1680  
Db 1950 AACACGCGCGCCCTTACAGGCAAGTGTGCTGCTGTATCATGATGATGATGATGATGATGATG 2009  
Qy 1681 ACCAAGCTGTGGGGGCGCCCGCTGCAATCTGGGGGGGCTGGGCAACAACATTTGATC 1740  
Db 2010 ACCAAGCTGTGGGGGCGCCCGCTGCAATCTGGGGGGGCTGGGCAACAACATTTGATC 2069  
Qy 1741 TGCCCCAGGAGCTTCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTGGGG 1800  
Db 2070 TGCCCCAGGAGCTTCCGGAAGCATCCGAGGCACTTACACCAATGCGGTTGGGG 2129  
Qy 1801 CTTGGTTG 1809  
Db 2130 CCGTGGTTG 2138

RESULT 9  
US-08-612-973-49  
; Sequence 49, Application US/08612973  
; Patent No. 6150134  
; GENERAL INFORMATION:  
; APPLICANT: MAERTENS, GEBERT  
; APPLICANT: BOSMAN, FONS  
; APPLICANT: DE MARYNOFF, GUY  
; APPLICANT: BUYSSE, MARIE-ANGE  
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE  
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHAYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,973

```

FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURES:
NAME/KEY: CDS
LOCATION: 1..2430
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2427
US-08-612-973-49:

Query Match      85.1%; Score 1538.6; DB 3; Length 2433;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

1  ATGAGCAGCAATCTTAACCTCAAGAAAACCAACGATACCAACCGCCGCCACAG 60
1  ATGAGCAGCAATCTTAACCTCAAGAAAACCAACGATACCAACCGCCGCCACAG 60
61  GACGTCAAGTCCCGGCGGTGTGATGATGATGATGATGATGATGATGATGATGAT 120
61  GACGTCAAGTCCCGGCGGTGTGATGATGATGATGATGATGATGATGATGATGAT 120
121  GGGCCCAAGTGGGTGTGCGCGGATGATGATGATGATGATGATGATGATGATGAT 180
121  GGGCCCAAGTGGGTGTGCGCGGATGATGATGATGATGATGATGATGATGATGAT 180
181  AGGCGCAACCTATCCCAAGGCTGCGCAAGCCGAGGCGGCTGCGTCAAGCCCGG 240
181  AGGCGCAACCTATCCCAAGGCTGCGCAAGCCGAGGCGGCTGCGTCAAGCCCGG 240
241  TACCTTGGCCCTCTATGAGCAAGAGGATGAGGATGAGGATGAGGATGAGGATGAG 300
241  TACCTTGGCCCTCTATGAGCAAGAGGATGAGGATGAGGATGAGGATGAGGATGAG 300
301  CGGCGCTCGGCTGATGAGGAGGCTTACAGACCCCGGCTGAGGCTGATGAGGAT 360
301  CGGCGCTCGGCTGATGAGGAGGCTTACAGACCCCGGCTGAGGCTGATGAGGAT 360
361  AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
361  AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
421  GGGCGCTCCCTTAAGGAGGCGCTGCAAGGCGCTGCGCATGAGCGTCCGAGTTC 480
421  GGGCGCTCCCTTAAGGAGGCGCTGCAAGGCGCTGCGCATGAGCGTCCGAGTTC 480
481  GGGCGCTCCCTTAAGGAGGCGCTGCAAGGCGCTGCGCATGAGCGTCCGAGTTC 540
481  GGGCGCTCCCTTAAGGAGGCGCTGCAAGGCGCTGCGCATGAGCGTCCGAGTTC 540
541  TGGCGTCCCTTAAGGAGGCGCTGCAAGGCGCTGCGCATGAGCGTCCGAGTTC 600
541  TGGCGTCCCTTAAGGAGGCGCTGCAAGGCGCTGCGCATGAGCGTCCGAGTTC 600
601  TACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
601  TACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

```

```

661  ATGCAACACCCCGGAGTGTGACCTGTGTCGAGGAGGATTAATTCCTCCGCTGTGAGTA 720
661  ATGCAACACCCCGGAGTGTGACCTGTGTCGAGGAGGATTAATTCCTCCGCTGTGAGTA 720
721  GGGCTCACTCCACAGCTGCGGCGCAAGAGCGCAGCATCCCACTGCGCAAAATACAGCG 780
721  GGGCTCACTCCACAGCTGCGGCGCAAGAGCGCAGCATCCCACTGCGCAAAATACAGCG 780
781  CAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
781  CAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
841  TGGGATCTGTTTCTCTGCTCTGACGATGATGATGATGATGATGATGATGATGATGAT 900
841  TGGGATCTGTTTCTCTGCTCTGACGATGATGATGATGATGATGATGATGATGATGAT 900
901  GTACAGAGCTGCAATGTTCAATCTATCCCGGCAAGTATCAGTCAACGCGATGCTTGG 960
901  GTACAGAGCTGCAATGTTCAATCTATCCCGGCAAGTATCAGTCAACGCGATGCTTGG 960
961  GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
961  GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1021  CCAAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1021  CCAAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1081  TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1081  TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1141  GTTGAAGGGAACCTTACAGACAGAGGAGGACACAGCGCGCGCCCAAGGCTTACA 1200
1141  GTTGAAGGGAACCTTACAGACAGAGGAGGACACAGCGCGCGCCCAAGGCTTACA 1200
1201  TCCCTCTTACACCTGAGGCGGCTGCAAAATTCAGCTTGTAAACCAACGAGCGCTGG 1260
1201  TCCCTCTTACACCTGAGGCGGCTGCAAAATTCAGCTTGTAAACCAACGAGCGCTGG 1260
1261  CACATCAACAGAACTGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1320
1261  CACATCAACAGAACTGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1320
1321  CTATTCCTAACAACAATTAATCACTGCTGATGATGATGATGATGATGATGATGATGAT 1380
1321  CTATTCCTAACAACAATTAATCACTGCTGATGATGATGATGATGATGATGATGATGAT 1380
1381  CCCATTGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1381  CCCATTGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1441  TCCATTCGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1441  TCCATTCGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1501  GAGGTGTGAGCCCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1501  GAGGTGTGAGCCCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1561  CGTTTGGGCGCCCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1561  CGTTTGGGCGCCCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1621  AACACGCGGCGCGCAAGGAGCAATGATGATGATGATGATGATGATGATGATGATGAT 1680
1621  AACACGCGGCGCGCAAGGAGCAATGATGATGATGATGATGATGATGATGATGATGAT 1680
1681  ACCAAGAGCTGTGAGGAGGCGCCCGTGAACATGAGGAGGAGTGGCAACAACCTTATG 1740
1681  ACCAAGAGCTGTGAGGAGGCGCCCGTGAACATGAGGAGGAGTGGCAACAACCTTATG 1740

```





Db 1141 GTGACAGGCGATACCGCGGTGTGACAGAGGGGACAGACCTCCGATACAGAGGGGCTTTG16 1200  
 QY 1201 TCCCTCTTACACCTTGGGCGGCTCAGAAATTCACCTTGTAAACCAACGGCAGCTTG 1260  
 Db 1201 TCCCTCTTAAAGCCCGGGTGGGTGAGAAATTCACCTTGTAAACCAACGGCAGCTTG 1260  
 QY 1261 CACATCAAGAACTGCTTGAATGCAATGATCTCCCTCAAACTGGGTTCTTTGGCGG 1320  
 Db 1261 CACATCAAGAACTGCTTGAATGCAATGATCTCCCTCAAACTGGGTTCTTTGGCGG 1320  
 QY 1321 CTGTTCTACACGACAGGTTCAATGCGTCCGATGCTCAGAGGCACTGGCAGCTGCGC 1380  
 Db 1321 CTATCTTCAAAACAAATTCACCTGCTGATGCTCCAGAGGCTTGCGCAGCTGCGC 1380  
 QY 1381 CCCATTGACAGCTTGCATCAGGGGTGGGTTCCATCACTTAATGATGCCACGCTTG 1440  
 Db 1381 TCCATCGAAGATTCCTCAGAGGGTGGGTTCCCTCACTTAACCTAGCCTAACAGCTG 1440  
 QY 1441 GACAGAGGCGCTTATTTGCTGACCTACGACCTCAACCGGTGTATGCGCCGCGGTTG 1500  
 Db 1441 GACAGAGGCGCTTATTTGCTGACCTACGACCTCAACCGGTGTATGCGCCGCGGTTG 1500  
 QY 1501 CAGGTGTGGCCCACTGTATCTGTTTCACTCCAAAGCCTGTTGTGGTGGGAGACGAT 1560  
 Db 1501 CAGGTGTGGCCCACTGTATCTGTTTCACTCCAAAGCCTGTTGTGGTGGGAGACGAT 1560  
 QY 1561 CGTTTGGGCGCCCTTACGATGAGATGGGTGAAATGAGACGAGCTGCTGCTTCTAAC 1620  
 Db 1561 CGGTTTGGGTGCTCCCACTGATTAATGAGGGGGGAGACGAGCTGATGTGCTGATTCAC 1620  
 QY 1621 AACAGCGGCGCGGACAGGGGCACTGATTCGCTGTATCAATGATGATGACCGGGGTTG 1680  
 Db 1621 AACAGCGGCGCGGACAGGGGCACTGATTCGCTGTATCAATGATGATGACCGGGGTTG 1680  
 QY 1681 ACCAAGACGTGTGGGGGCGCCCGCTGCAATCGGGGGGGTCCGCAACAACCTTTGATC 1740  
 Db 1681 ACCAAGACGTGTGGGGGCGCCCGCTGCAATCGGGGGGGTCCGCAACAACCTTTGATC 1740  
 QY 1741 TGGCCCAAGGAGCTGCTTCCGGAAGCATCCGAGGCACTTAACCAATGCGGTTGGGG 1800  
 Db 1741 TGGCCCAAGGAGCTGCTTCCGGAAGCATCCGAGGCACTTAACCAATGCGGTTGGGG 1800  
 QY 1801 CCTTGGTTG 1809  
 Db 1801 CCTTGGTTG 1809

## RESULT 11

US-08-150-204E-96  
 Sequence 96, Application US/08150204E  
 Patent No. 6538126

## GENERAL INFORMATION:

APPLICANT: CHO, Joong Myung  
 LEE, Young Beom  
 PARK, Young Moo  
 LIM, Kook Jin  
 CHOI, Deog Young  
 SO, Hong Seob  
 KIM, Chun Hyung  
 KIM, Sung Taek  
 YANG, Jae Young  
 TITLE OF INVENTION: HEPATITIS C DIAGNOSTICS AND VACCINES  
 NUMBER OF SEQUENCES: 128  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: YANG, Jae Young  
 STREET: 386-1, Doryong-dong, Yuseong-gu  
 CITY: Daejeon  
 STATE: Daejeon  
 COUNTRY: Republic of Korea  
 ZIP: 305-340

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5inch, 1.44MB storage  
 COMPUTER: IBM PC/pentium

OPERATING SYSTEM: Windows  
 SOFTWARE: Microsoft Word  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/150,204E  
 FILING DATE: 20-Apr-1994  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: KR 91-9510  
 FILING DATE: 10-JUN-1991  
 APPLICATION NUMBER: KR 91-13601  
 FILING DATE: 6-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Shahan Islam, Esq.  
 REGISTRATION NUMBER: 32,507  
 REFERENCE/DOCKET NUMBER: 2695/FLK  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 940-8564  
 TELEFAX: (212) 940-8776  
 INFORMATION FOR SEQ ID NO: 96  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9472 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA  
 FEATURE:  
 OTHER INFORMATION: KHCV-LBC1, Fig. 2  
 SEQUENCE DESCRIPTION: SEQ ID NO: 96  
 US-08-150-204E-96

Query Match 84.6%; Score 1531.2; DB 4; Length 9472;  
 Best Local Similarity 90.4%; Pred. No. 0;  
 Matches 1655; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 1 ATGAGCAGCAATCTTAAACCTCAAGAAACCAACGTAACCAACCGCCGCCACAG 60  
 Db 343 ATGAGCAGCAATCTTAAACCTCAAGAAACCAACGTAACCAACCGCCGCCACAG 402  
 QY 61 GACGTCAAGTCCCGGCGGTGTGATGATGCTTGTGAGATTACTGTGGCGGCAAG 120  
 Db 403 GATATTAATTTCCCGGCGGTGTGATGATGCTTGTGAGATTACTGTGGCGGCAAG 462  
 QY 121 GGGCCAGGTTGGGTGTGGCGGAGCTAGAAAGACTTCCGAGGGTGGCACTCGTGA 180  
 Db 463 GGGCCAGGTTGGGTGTGGCGGAGCTAGAAAGACTTCCGAGGGTGGCACTCGTGA 522  
 QY 463 GGGCCAGGTTGGGTGTGGCGGAGCTAGAAAGACTTCCGAGGGTGGCACTCGTGA 522  
 Db 181 AGGCGCAACCTATCCCAAGGCTCGCCAGCCCGAGGGGAGGAGGCTGAGCCCGG 240  
 QY 523 AGGCGCAACCTATCCCAAGGCTCGCCAGCCCGAGGGGAGGAGGCTGAGCCCGG 582  
 QY 241 TACCTTGGGCGCTTATGGAACAGAGGCAATGGGGTGGGAGATGGCTCTGTACCC 300  
 Db 583 TACCTTGGGCGCTTATGGAACAGAGGCAATGGGGTGGGAGATGGCTCTGTACCC 642  
 QY 301 CGTGGCTCCGGGCTAGTTGGGGGCCCACTGACCCCGGCGTAAAGTGGCTAATTGGGT 360  
 Db 643 CGTGGCTCCGGGCTAGTTGGGGGCCCACTGACCCCGGCGTAAAGTGGCTAATTGGGT 702  
 QY 643 CGTGGCTCCGGGCTAGTTGGGGGCCCACTGACCCCGGCGTAAAGTGGCTAATTGGGT 702  
 Db 361 AAAGTCATGATACCTTCATGCGGCTTGGCGGACCTTCATGGGGTACATTCGCTGTC 420  
 QY 703 AAAGTCATGATACCTTCATGCGGCTTGGCGGACCTTCATGGGGTACATTCGCTGTC 762  
 Db 421 GGGGCTCCCTTGGGGGCGCTGCGAGGCGCTTGGCGGACCTTCATGGGGTACATTCGCTGTC 480  
 QY 763 GGGGCTCCCTTGGGGGCGCTGCGAGGCGCTTGGCGGACCTTCATGGGGTACATTCGCTGTC 822  
 Db 481 GGGGCTCCCTTGGGGGCGCTGCGAGGCGCTTGGCGGACCTTCATGGGGTACATTCGCTGTC 540  
 QY 823 GGGGCTCCCTTGGGGGCGCTGCGAGGCGCTTGGCGGACCTTCATGGGGTACATTCGCTGTC 882  
 Db 541 TTGCTGTCTGTTTGAACCAATTCAGCTTCCGTTATGAAGTCCGCAACGCTGCCGATC 600  
 QY 883 CTGCTGTCTGTTTGAACCAATTCAGCTTCCGTTATGAAGTCCGCAACGCTGCCGATC 942





TOPOLOGY: linear  
MOLECULE TYPE: cDNA from genomic RNA  
FEATURES:  
NAME/KEY: misc\_feature  
LOCATION: 1..6039  
OTHER INFORMATION: /note: "sequence = 333 - 6371 of  
OTHER INFORMATION: SEQ ID NO: 1"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..6039  
US-08-324-977-11

Query Match 83.9%; Score 1517.8; DB 1; Length 6039;  
Best Local Similarity 89.9%; Pred. No. 0;  
Matches 1627; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 1 ATGAGCAGCATCTTAACCTCAAGAAAACCAAGTAACACCAACCGCCGCCACAG 60  
DB 1 ATGAGCAGCATCTTAACCTCAAGAAAACCAAGTAACACCAACCGCCGCCACAG 60

QY 61 GAAGTCAGATCCCGGCGGCTGTGATGATGCTGTGAGATTACTGTGCGCGCAGG 120  
DB 61 GAAGTCAGATCCCGGCGGCTGTGATGATGCTGTGAGATTACTGTGCGCGCAGG 120

QY 121 GGGCCAGGTTGGGTGGCGCGCATAGGAAGACTTCGAGGGGTGCAACCTCGTGA 180  
DB 121 GGGCCAGGTTGGGTGGCGCGCATAGGAAGACTTCGAGGGGTGCAACCTCGTGA 180

QY 181 AGCGCAGCAACTTATCCCAAGGCTCGCAGCCGAGGGCAGGGCTCGAGCCGAG 240  
DB 181 AGCGCAGCAACTTATCCCAAGGCTCGCAGCCGAGGGCAGGGCTCGAGCCGAG 240

QY 241 TACCTTTGGCCCTCTATAGGCAAGAGGATGAGGATGAGGATGAGGATGAGG 300  
DB 241 TACCTTTGGCCCTCTATAGGCAAGAGGATGAGGATGAGGATGAGGATGAGG 300

QY 301 CGGAGCTCCGCGCTAGTGGGGCCCACTGACCCCGGCGTGGTGGCTAATTGGGT 360  
DB 301 CGGAGCTCCGCGCTAGTGGGGCCCACTGACCCCGGCGTGGTGGCTAATTGGGT 360

QY 361 AAGATCATGATACCTCATAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 420  
DB 361 AAGATCATGATACCTCATAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 420

QY 421 GGGCGCTCCCTTAAAGGGGCGCTGCAAGGCGCTGAGGCGATGAGGCTGAGG 480  
DB 421 GGGCGCTCCCTTAAAGGGGCGCTGCAAGGCGCTGAGGCGATGAGGCTGAGG 480

QY 481 GGGGTGAACATGCAACAGGGAATTTAACCGGCTGCTCTATCTTCTCTGGCT 540  
DB 481 GGGGTGAACATGCAACAGGGAATTTAACCGGCTGCTCTATCTTCTCTGGCT 540

QY 541 TTGCTGTCTGTTGACATTCAGCTTCCGTTATGAGTGGCAACGTCGCGGATC 600  
DB 541 TTGCTGTCTGTTGACATTCAGCTTCCGTTATGAGTGGCAACGTCGCGGATC 600

QY 601 TACCATGTACGAAAGATGCTCAACTCAAGATCGTGTACGAGACAGGGAATG 660  
DB 601 TACCATGTACGAAAGATGCTCAACTCAAGATCGTGTACGAGACAGGGAATG 660

QY 661 ATGCAACACCCCGGGGTGTGCGCTGTGTCGAGGAGGTAATCTCTCCGCTGT 720  
DB 661 ATGCAACACCCCGGGGTGTGCGCTGTGTCGAGGAGGTAATCTCTCCGCTGT 720

QY 721 GGGCTCACTCCAGCGCTCGAGGCAAGGATGATCACTCCCAACAGAGATACAG 780  
DB 721 GGGCTCACTCCAGCGCTCGAGGCAAGGATGATCACTCCCAACAGAGATACAG 780

QY 781 CAGCTGATGATGCTGCTGGGGCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 781 CAGCTGATGATGCTGCTGGGGCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCT 840

QY 841 TGGGATCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

DB 841 TGGGATCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

QY 901 GTACAGAGCTGCAATGTTCAATCTATCCCGGCAAGTACAGGTACAGGATGCTGG 960  
DB 901 GTACAGAGCTGCAATGTTCAATCTATCCCGGCAAGTACAGGTACAGGATGCTGG 960

QY 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
DB 961 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020

QY 1021 CCACAGCTGTGCTGAGCATGAGGCGGGGCGCACTGAGGAGTCTGCGCGGCTTGG 1080  
DB 1021 CCACAGCTGTGCTGAGCATGAGGCGGGGCGCACTGAGGAGTCTGCGCGGCTTGG 1080

QY 1081 TACTATTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
DB 1081 TACTATTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140

QY 1141 GTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
DB 1141 GTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200

QY 1201 TCCCTCTTCAACCTGAGCGGCTCAGAAAATTCAGCTTGTAAACACCAACGAGCTGG 1260  
DB 1201 TCCCTCTTCAACCTGAGCGGCTCAGAAAATTCAGCTTGTAAACACCAACGAGCTGG 1260

QY 1261 CACATCAACAGGAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
DB 1261 CACATCAACAGGAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320

QY 1321 CTGTTCTTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
DB 1321 CTGTTCTTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380

QY 1381 CCCATGACCAAGTTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
DB 1381 CCCATGACCAAGTTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

QY 1441 GACAGAGGAGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500  
DB 1441 GACAGAGGAGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500

QY 1501 CAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
DB 1501 CAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560

QY 1561 CGTTTGGGCGCCCTTACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
DB 1561 CGTTTGGGCGCCCTTACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620

QY 1621 AACACAGGAGCGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680  
DB 1621 AACACAGGAGCGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680

QY 1681 ACCAAGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740  
DB 1681 ACCAAGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740

QY 1741 TGGCCAGAGGAGCTGCTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800  
DB 1741 TGGCCAGAGGAGCTGCTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800

QY 1801 CCTTGGTTG 1809  
DB 1801 CCTTGGCTG 1809

RESULT 13  
US-08-384-616-11  
; Sequence 11, Application US/08384616  
; Patent No. 5847101  
; GENERAL INFORMATION:

APPLICANT: OKAYAMA, Hiroto  
APPLICANT: FUKU, Isao  
APPLICANT: MORI, Chisato  
APPLICANT: TAKAMIZAWA, Akahisa  
APPLICANT: YOSHIDA, Iwao  
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
NUMBER OF SEQUENCES: 50  
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Armstrong, Westerman, Hattori, Mcleland &  
STREET: 1725 K St. N.W. Suite 1000  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/384,616  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/769,996  
FILING DATE: 02-OCT-1991  
APPLICATION NUMBER: JP 2-167466  
FILING DATE: 25-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-230921  
FILING DATE: 31-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-305605  
FILING DATE: 09-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/635,451  
FILING DATE: 28-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens-Smith, Theresa M.  
REGISTRATION NUMBER: 36,281  
REFERENCE/DOCKET NUMBER: 900703B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2930  
TELEFAX: (202) 887-0357  
TELEX: 440142  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6039 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULAR TYPE: cDNA from genomic RNA  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..6039  
OTHER INFORMATION: /note: "sequence = 333 - 6371 of  
OTHER INFORMATION: SEQ ID NO: 1"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..6039  
US-08-384-616-11

Query Match 83.9%; Score 1517.8; DB 2; Length 6039;  
Best Local Similarity 89.9%; Pred. No. 0;  
Matches 1627; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 1 ATGAGCAGATCTCTAAACCTCAAGAAACCAACGTAACCAACCGCCGCCACG 60  
DB 1 ATGAGCAGATCTCTAAACCTCAAGAAACCAACGTAACCAACCGCCGCCACG 60  
QY 61 GACGTCAAGTCCCGGCGGTGTCAGATCGTTGTGAGATTACCTGTTGCCGCGCAGG 120

DB 61 GACGTCAAGTCCCGGCGGTGTCAGATCGTTGTGAGATTACCTGTTGCCGCGCAGG 120  
QY 121 GCGCCAGGTGGGTGTGGCGGCGCATAGAAAGCTTCCGACCGCTCCACCTCTGGA 180  
DB 121 GCGCCAGGTGGGTGTGGCGGCGCATAGAAAGCTTCCGACCGCTCCACCTCTGGA 180  
QY 181 AGGCGACAACCTATCCCAAGGCTGCGCAGCCCAAGGCGAGGCTGAGCCCGG 240  
DB 181 AGGCGACAACCTATCCCAAGGCTGCGCAGCCCAAGGCGAGGCTGAGCCCGG 240  
QY 241 TACCCCTGGCCCTCTATAGCGACGAGGCGATGAGGTGGGAGAGATGCTCTGACCC 300  
DB 241 TACCCCTGGCCCTCTATAGCGAATGAGGCTTATAGGTGGGAGAGATGCTCTGACCC 300  
QY 301 CTGTGCTCCCGGCTAGTGGGGGCGCCCATACCCCGGCGTAGTGGCGTAATTGGGT 360  
DB 301 CTGTGCTCCCGGCTAGTGGGGGCGCCCATACCCCGGCGTAGTGGCGTAATTGGGT 360  
QY 361 AAAGTCATCATACCTCTACATGCGGCTTCCGCACTCATGGGGTACATTCCGCTGTC 420  
DB 361 AAAGTCATCATACCTCTACATGCGGCTTCCGCACTCATGGGGTACATTCCGCTGTC 420  
QY 421 GCGCTCCCTTAAAGGGGCGCTGCGAGGGCCCTGCGCATGGCGTCCGGCTTCTGAGGAC 480  
DB 421 GCGCTCCCTTAAAGGGGCGCTGCGAGGGCCCTGCGCATGGCGTCCGGCTTCTGAGGAC 480  
QY 481 GCGCGTAATATGACAAGGGAATTTACCGGTTGCTCTTCTATCTTCTCTGCTGCT 540  
DB 481 GCGCGTAATATGACAAGGGAATTTACCGGTTGCTCTTCTATCTTCTCTGCTGCT 540  
QY 541 TTGCTGTCTGTTTACCATTCAGCTTCCGCTTATGAGTGGCGCAAGTGTCCGGATC 600  
DB 541 TTGCTGTCTGTTTACCATTCAGCTTCCGCTTATGAGTGGCGCAAGTGTCCGGATC 600  
QY 601 TACCATGTCAAGAACGATGCTTCCATCACTCAAGCATCTGTATGAGACAGCGAATGATC 660  
DB 601 TACCATGTCAAGAACGATGCTTCCATCACTCAAGCATCTGTATGAGACAGCGAATGATC 660  
QY 661 ATGCAACACCCCGGGGTGTGGTCCCTGTGTCCGGGAGGTAATCTCCCGGTGCTGGTA 720  
DB 661 ATGCAACACCCCGGGGTGTGGTCCCTGTGTCCGGGAGGTAATCTCCCGGTGCTGGTA 720  
QY 721 GCGCTCACTCCACGCTGCGGCGCAAGGACGCGACATCCCACTGCGCAATATGACGCG 780  
DB 721 GCGCTCACTCCACGCTGCGGCGCAAGGACGCGACATCCCACTGCGCAATATGACGCG 780  
QY 781 CACGTGATTTGCTGTGGGGGCTGCTTGTGCTTCCGCTATGATGATGATGATGATGATC 840  
DB 781 CACGTGATTTGCTGTGGGGGCTGCTTGTGCTTCCGCTATGATGATGATGATGATGATC 840  
QY 841 TGGCGATCTGTTTCTCTGCTCTCAAGCTGTTCACTTCCGCTCCGCGCATCATACG 900  
DB 841 TGGCGATCTGTTTCTCTGCTCTCAAGCTGTTCACTTCCGCTCCGCGCATCATACG 900  
QY 901 GTACAGGACTCAATTTGTTCAATCTATCCCGGCAAGATCAGGTACACCGGATGCTTGG 960  
DB 901 GTACAGGACTCAATTTGTTCAATCTATCCCGGCAAGATCAGGTACACCGGATGCTTGG 960  
QY 961 GATATGATGATGAACCTGTCACTTACAGACAGCCCTAGTGTATGCACTACTCCGATC 1020  
DB 961 GATATGATGATGAACCTGTCACTTACAGACAGCCCTAGTGTATGCACTACTCCGATC 1020  
QY 1021 CCACAGCTGTGCGGACATGATGATGCGGGGCGGCGGAGTCCGCGGCGCTTGGC 1080  
DB 1021 CCACAGCTGTGCGGACATGATGATGCGGGGCGGCGGAGTCCGCGGCGCTTGGC 1080  
QY 1081 TACTATTTCATGTGGGGAATGAGGTCTTGTGTTGATGATGATGATGATGATGATGATC 1140  
DB 1081 TACTATTTCATGTGGGGAATGAGGTCTTGTGTTGATGATGATGATGATGATGATGATC 1140  
QY 1141 GTTGAACGGGAACCTTACACGACAGGAGGAGACACAGGCGCGCGCCCAAGGCTTACA 1200  
DB 1141 GTTGAACGGGAACCTTACACGACAGGAGGAGACACAGGCGCGCGCCCAAGGCTTACA 1200

Db	1141	GTGTACGGGGATACCACAGTGAACAGGGGGGGGGCAAGCCAAACCAACACAGGCTGGT	1200
QY	1201	TCCTCTTTCACACTCTGGGCGCGCTTCAGAAAATCCAGCTTTGTAACCAACGCACTGG	1266
Db	1201	TCCAGTTTCGCAAGGGGGCGCTCTCAGAAAAATCCAGCTTTAATAACCAACCAAGGGAGTTGG	1260
QY	1261	CACATCAACAGAACTGCTTGAACTGGCAATGACTCCCTCCAAACTGGGGTTCTTGGCGCG	1350
Db	1261	CACATCAACAGAACTGCTTGAACCTGCAATGACTCTCTCCAGACTGGGGTTCTTGGCGCG	1320
QY	1321	CTGTTCTTACACGCACAGGTTCAATGCGTCCGGATGCTCAGAGCGCATGGCAGACTCCGC	1380
Db	1321	CTGTTCTTACACACTAGTTTCAACTCGTCCGGGTGCCAGAGCGCATGGCAGACTGGCGCG	1380
QY	1381	CCCAATTGACCAAGTTGCGATCAGGGGGGGGGTCCCACTATTAATAGATCCCAAGCGTGG	1440
Db	1381	ACCATATGCAAGTTGCGACCAAGGAGTGGGGTCCCACTTATATGCTGAATCTAGACAGATCA	1440
QY	1441	GACCAAGAGCCCTATTGCTGGCACTACCACTCAACCGTGTGGATGTCGCCGCGTTG	1500
Db	1441	GACCAAGAGCCCATATTGTGGCACTACCACTCAACCGTGTGGATGTCGCCGCGTTG	1500
QY	1501	CAGGTGTGTGGCCCAAGTGACTGTTTCACTCCAAAGCCTGTTGTGTGGGGACAACGAT	1566
Db	1501	GAGGTGTGTGGCCCAAGTGACTGCTTCAACCCCAAGCCTGTGCTGTGGGGACAACCGAT	1560
QY	1561	CGTTTCCGCGCCCTCAACGTACAGATGGGGGTGAGAAATGAGACGGAACGTCTCTTCTCAAC	1620
Db	1561	CGTTTCCGCTGTCCTCAAGTAATAGATGGGGGGAGAAACGAGACTGACTCTCTCTCAAC	1620
QY	1621	AACACGCGGCGGCCACGGGGACAATGTTGCGCTGTACATGATGATGAATAGCAACGGGTTT	1680
Db	1621	AACACGCGGCGGCCGCAAGGCAACTGGTTTCCGCTGTACATGATGATGAATAGCAACGGGTTT	1680
QY	1681	ACCAAGACGTGTGGGGGGCCCCCGGTGCAACTCGGGGGGGTTCGGCAACAACATTTGATC	1740
Db	1681	ACCAAGACATGTGGGGGGCCCCCGGTGAACATCGGGGGGGTTCGGCAACAACCTTGATC	1740
QY	1741	TGCCCCACGACTGCTTTCGGAAGCATCCAGGCACTTACACCAATGTGGTTGGGG	1800
Db	1741	TGCCCCACGACTGCTTTCGGAAGCAACCCGAGGCTTACACCAAAATGTGGTTGGGG	1800
QY	1801	CTTTGGTTG 1809	
Db	1801	CTTTGGCTG 1809	

RESULT 14  
 US-08-904-686A-11  
 Sequence 11, Application US/08904686A  
 Patent No. 5998130  
 GENERAL INFORMATION:  
 APPLICANT: OKAYAMA, Hiroto  
 APPLICANT: FUKU, Ieao  
 APPLICANT: MORI, Chisato  
 APPLICANT: TAKAMIZAWA, Akahisa  
 APPLICANT: YOSHIDA, Iwao  
 TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
 TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
 NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSSE: Armstrong, Westernm, Hattori, Mclelland &  
 STREET: 1725 K St., N.W. Suite 1000  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 in., 1.44MB  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 SOFTWARE: ASCII

```

1 CURRENTAPPLICATION DATA:
2 APPLICATION NUMBER: US/08/904,686A
3 FILING DATE: 01-AUG-1997
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: US 08/324,977
6 FILING DATE: 18-OCT-1994
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: JP 2-167466
9 FILING DATE: 25-JUN-1990
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: JP 2-230921
12 FILING DATE: 31-AUG-1990
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: JP 2-305605
15 FILING DATE: 09-NOV-1990
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US 08/099,706
18 FILING DATE: 30-JUL-1993
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: US 07/769,996
21 FILING DATE: 02-OCT-1991
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US 07/635,451
24 FILING DATE: 28-DEC-1990
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Mclelland, Le-Nhung
27 REGISTRATION NUMBER: 31,541
28 REFERENCE/DOCKET NUMBER: 900703G
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (202) 659-2930
31 TELEFAX: (202) 867-0357
32 INFORMATION FOR SEQ ID NO: 11:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 6039 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: single
37 TOPOLOGY: linear
38 MOLECULE TYPE: cDNA from genomic RNA
39 FEATURE:
40 NAME/KEY: misc_feature
41 LOCATION: 1..6039
42 OTHER INFORMATION: /note: "sequence = 333 - 6371 of
43 OTHER INFORMATION: SEQ ID NO: 1"
44 FEATURE:
45 NAME/KEY: CDS
46 LOCATION: 1..6039
47 US-08-904-686A-11

```

Query Match	83.9%;	Score 1517.8;	DB 2;	Length 6039;
Best Local Similarity	89.9%;	Pred. No. 0;	Mismatches 182;	Indels 0;
Matches 1627;	Conservative	0;	Gaps	0
Qy	1	ATGAGACGAACTCTTAAACCTCAAGAAAAACCAAGCTAACACCAACCGCGCCACAG	60	
Db	1	ATAGGACGAACTCTTAAACCTTAAAGAAAAACCAACCTTACACCAACCGCGCCACAG	60	
Qy	61	GACGTCAAGTCCCGGCGGCGGTGTCAGATCGTTGGAGTTTACTGTTGCGCGCAGG	120	
Db	61	GAGTCAAGTCCCGGCGGCGGTGTCAGATCGTTGGAGTTTACTGTTGCGCGCAGG	120	
Qy	121	GSCCCAGGTTGGGTGTGTGCGCGCACTAGAAAGCTTCGAGCGGTGCAACTCTGTGA	180	
Db	121	GSCCCAGGTTGGGTGTGTGCGCGCACTAGAAAGCTTCGAGCGGTGCAACTCTGTGA	180	
Qy	181	AGGCGCAACCTATCCCAAGGCTGCGACGCCAGGCGAGGGCTGAGGCTCAAGCCCGG	240	
Db	181	AGGCGCAACCTATCCCAAGGCTGCGACGCCAGGCGAGGAGCTTGGGCTCAAGCCCGG	240	
Qy	241	TACCTCTGGACCCCTCTATGACAAACAGAGGCAATGGGGTGTGGCAGANTGGTCTGTCA	300	
Db	241	TACCTCTGGACCTCTCTATGCAATAGGAGCTTTAGGGTGTGGCAGANTGGTCTGTCA	300	
Qy	301	CGTGCTCCCGACCTAGTTGGGGCCCACTGACCCCGCGGTAGGTGCGCTAAATTGGAT	360	

Db	301	CGGGCTCCGGCTAGTGTGGGGCCCAAGAACCCCGCGTAGGTCCGCTAATTTGGGT	360
Oy	361	AAAGTCATGATACCTTCACATATGCGGCTTGCCGCACTTCATGGGGTACATTTCCGTCGT	420
Db	361	AAGTCATGATACCTTCACATATGCGGCTTGCCGCACTTCATGGGGTACATTTCCGTCGT	420
Oy	421	GGCGTCCCTTAGGGGGCGCTGCGAGGGCCCTGGGCGATAGCGTCCGGGTTTGGAGAC	480
Db	421	GGCGCCCCCTGGGGGGCGCTGCCAGGGCCCTGGGCAATGTGTCCGGGTTCTGGAGAC	480
Oy	481	GAGCGAAGTATGCAACAGGGAAATTAACCGGTATGCTTTCTATCTTCTTGCT	540
Db	481	GAGCGAAGTATGCAACAGGGAAATGCGCGGTGTCTCTTCTATCTTCTTGCT	540
Oy	541	TTGCTGTCTGTGTTGACCAATTCACGCTTCCGCTTAATGAAATGGCGCAAGTGTCCGGATC	600
Db	541	CTGCTGTCTGTGCTGACCAACCCACGCTTCCGCTTAATGAAATGGCGCAAGTGTCCGGATC	600
Oy	601	TACCATGTACAGAACGATTTGCTCCACTTCAGCATGTGTACGAGACAGCGGACATGATC	660
Db	601	TATCATGTACAGAACGACTGTCCAAACGCAATGTGTATGAGGACAGCGGACTTGATC	660
Oy	661	ATGCACAACCCCGGGGTGTGCCGTGTGTCGGGAGGGTAATTCCTCCGCTGCTGGATA	720
Db	661	ATGCATACCTCCGGGTGTGCCGTGTGTCGGGAGGCAATCTCTCCGCTGCTGGATA	720
Oy	721	GCGCTCACTCCACGCTGCGCGGCAAGGACGCGACAGATCCCACTGCGACATAAGACGC	780
Db	721	GCGCTCACTCCACGCGCTGCGAGCGAGAACGTCAACATCCCAACAGACATAGACGC	780
Oy	781	CACGTGCATTTGCTGCTGGGGCGGCTGCTTGTCTCCGCTATGATATGATGATGGGGATCTC	840
Db	781	CACGTGCATTTGCTGCTGGGGCGGCTGCTTGTCTCCGCTATGATATGATGATGGGGATCTC	840
Oy	841	TGCGGATCTGTTTCTCTGCTCTACGCTTCACTTCGCGCTGCGGACATAGAC	900
Db	841	TGCGGATCTGTTTCTCTGCTCTCTACGCTTCACTTCGCGCTGCGGATATGACA	900
Oy	901	GTAACAGACTGCAATTTGTTCAATCTATCCCGGCAAGTACAGTACCGCATGGCTTGG	960
Db	901	TTAAGAGACTGTAATGCTCAATTTATCCCGGCAATGTGTGGGGTCAACGATGGCTTGG	960
Oy	961	GATATGATGATGAATGTGTCATCAACAGACGCCATAGTGTATGCGAGCTACTCCGATC	1020
Db	961	GACATGATGATGAATGTGTCGCGCCCAACAGGCCCTAGTGTGTGCGAGTTACTCCGATC	1020
Oy	1021	CCACAAGCTGTGCTGTGACATGTGTGGCGGGGGCCACTGGGAGTCTTGGCGGGCTTGGC	1080
Db	1021	CCACAAGCGCTGTGTGACATGTGTGGCGGGGGCCACTGGGAGTCTTGGCGGGCTTGGC	1080
Oy	1081	TACTATTCATATGATGGGGAATGCGGCTTAAGGCTTGGTGTGATGCTACTCTTGGCGGAC	1140
Db	1081	TACTATTCATATGCGGGGAATGCGGCTTAAGGCTTGTGATTTGATGCTACTCTTGTCTGGAC	1140
Oy	1141	GTTGACGGGGAACTTTAACAGACAGGGGGGACACAGCGCGCGGCCGCCACGCGCTTCA	1200
Db	1141	GTTGACGGGGAAATCCCAAGTACAGGGGGGGGCAAGCAAAACACCAACAGGCTGTG	1200
Oy	1201	TCCCTCTTCAACCTGGGGCGGCTCAAGAAATCCAGCTTTGTAACACCAACGCGACTGG	1260
Db	1201	TCCATGTTGCAAGTGGGCGGCTTCAGAAATCCAGCTTATTAACCAATATGGGAGTTGG	1260
Oy	1261	CACATCAACAGAACTGCGCTGTAACATGATGCTTCTTCAGACTGGGGTTTCTTTCGCG	1320
Db	1261	CACATCAACAGGACTGCGCTGTAACATGATGCTTCTTCAGACTGGGGTTTCTTTCGCG	1320
Oy	1321	CTGTTCTTACACGACAGGTTCAATGCTGTCCGAGACTGACAGCGCATGAGCGGACGCTGC	1380
Db	1321	CTGTTCTTACACATATGTTTCAACTGTGTCCGAGTCCCAAGCGCATATGGCCCAAGTCCGC	1380
Oy	1381	CCCATGACCAATGATCAGGGGTGGGGTCCCATCATTTATATGAGTCCCAAGGCTTGG	1440

Db	1381	ACACTTGAACAAGTTTCGACACGAGGANTGGGGTCCCACTTAATTAGTGCAGACTACGAGATCA	14410
Oy	1441	GACCAAGAGCCCTATTGTGTGCACTAGCACTCAACCGTGTGATTCGTCGCCGCTTG	1500
Db	1441	GACCAAGAGCCATTTGTGTGCACTAACCACTTCACAAAATGACATGTACTGTGCGTCG	15000
Oy	1501	CAGGTGTGTGGCCCAGTGTACTGTTTCACTTCAAGCCCTGTTGTGTGTGGGACGACCGAT	1560
Db	1501	GAGGTGTGTGGCCCAGTGTACTGCTTCAACCCCAAGCCCTGTGCTGTGTGTGGGACGACGAT	15600
Oy	1561	CGTTTCGCGCCCTTACGTACAGATGTGGGTGTGAATAGACGAGACGTGCTCTCTCAAC	1620
Db	1561	CGTTTCGCTGTCCCTTAGATGTAGATGTGGGGGAGAACAGACTGACGTGTGCTCTCTCAAC	16200
Oy	1621	AACACGGGGCCGCCACGGGGGCAATTGGTTGCGGCTGTACATGGAATGAATAGACCGGGTTC	1680
Db	1621	AACACGGGGCCGCCACGGGCAACATGTGTTGCGCTGTACATGGAATGAATAGACCGGGTTC	16800
Oy	1681	ACCAAGAAGTGTGGGGGCCCCCGGTGCAACAATCGGGGGGTGCGCAACAACATTTGATC	1740
Db	1681	ACCAAGAATGTGGGGGGCCCCCGTGTAAACATCGGGGGGGTTCGGCAACAACCTTGACC	17400
Oy	1741	TGCCCAACGGAATGCTTCCGGAAACATTCGGAAGGCACTTACACCAATGCGTTTCGGGG	1800
Db	1741	TGCCCAACGGAATGCTTCCGGAAACACCCGAGGCTACTACACAAAATGTGTGTTGGGG	18000
Oy	1801	CCTTGTGTTG 1809	
Db	1801	CCTTGTGCTG 1809	

RESULT 15-850-11  
 US-09-315-850-11  
 : Sequence 11, Application US/09315850  
 : Patent No. 6217872  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: OKAYAMA, Hiroto  
 : APPLICANT: FUKU, Isao  
 : APPLICANT: MORI, Chisato  
 : APPLICANT: TAKAMIZAWA, Akahisa  
 : APPLICANT: YOSHIDA, Isao  
 : TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
 : TITLE OF INVENTION: CNA AND ANTIGEN POLYPEPTIDE  
 : NUMBER OF INVENTION: 50  
 : NUMBER OF SEQUENCES: 50  
 : CORRESPONDENCE ADDRESSES:  
 : ADDRESS: Armstrong, Westernman, Hatiroi, McIeland &  
 : ADDRESS: Naughton  
 : STREET: 1725 K St. N.W. Suite 1000  
 : CITY: Washington  
 : STATE: D.C.  
 : COUNTRY: U.S.A.  
 : ZIP: 20006  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette, 3.5 in., 1.44MB  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 : SOFTWARE: ASCII  
 :  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/315,850  
 : FILING DATE:  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/904,686  
 : FILING DATE: 01-AUG-1997  
 : APPLICATION NUMBER: US 08/324,977  
 : FILING DATE: 18-OCT-1994  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: JP 2-167466  
 : FILING DATE: 25-JUN-1990  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: JP 2-230921  
 : FILING DATE: 31-AUG-1990  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: JP 2-305605

FILING DATE: 09-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/099,706  
FILING DATE: 30-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/769,996  
FILING DATE: 02-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/635,451  
ATTORNEY/AGENT INFORMATION:  
NAME: McLeod, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 900703G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2930  
TELEFAX: (202) 887-0357  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6039 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA from genomic RNA  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..6039 /note: "sequence = 333 - 6371 of  
OTHER INFORMATION: SEQ ID NO: 1"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..6039  
US-09-315-850-11

Query Match 83.9%; Score 1517.8; DB 3; Length 6039;  
Best Local Similarity 89.9%; Pred. No. 0;  
Matches 1627; Conservative 0; Mismatches 182; Indels 0; Gaps 0;  
QY 1 ATGAGACGAACTCTTAACTTCAAGAAAAACCAACGTTACCAACCGCCGCCACAG 60  
DB 1 ATGAGACGAACTCTTAACTTCAAGAAAAACCAACGTTACCAACCGCCGCCACAG 60  
QY 61 GAGGTCAAGTCCCGGCGGCGGTGTCAGATCGTTGAGTTCACCTGTTGCCGCGCAG 120  
DB 61 GAGGTCAAGTCCCGGCGGCGGTGTCAGATCGTTGAGTTCACCTGTTGCCGCGCAG 120  
QY 121 GAGCCGAGGTGGGTGTGCGCGAGTAAAGACTTCCGAGCGGTGCAACCTCGTGA 180  
DB 121 GAGCCGAGGTGGGTGTGCGCGAGTAAAGACTTCCGAGCGGTGCAACCTCGTGA 180  
QY 181 AGGCGAACAATTATCCCAAGGCTGCGCAGCCGAGGGGCAAGGCTGAGCCCGGG 240  
DB 181 AGGCGAACAATTATCCCAAGGCTGCGCAGCCGAGGGGCAAGGCTGAGCCCGGG 240  
QY 241 TACCTTGGCCCCCTATGSCAACAGAGGATGGGGGAGAGATGGCTCCGTACACC 300  
DB 241 TACCTTGGCCCCCTATGSCAACAGAGGATGGGGGAGAGATGGCTCCGTACACC 300  
QY 301 CGTGGCTCCCGGCTATGTTGGGGCCCACTGACCCCGGCGTAACTCGGTAAATTTGG 360  
DB 301 CGTGGCTCCCGGCTATGTTGGGGCCCACTGACCCCGGCGTAACTCGGTAAATTTGG 360  
QY 361 AAAAGTATGATACCTCTACATGCGGCTTCCGCACTTATGGGGTATCATTCGCTGTC 420  
DB 361 AAAAGTATGATACCTCTACATGCGGCTTCCGCACTTATGGGGTATCATTCGCTGTC 420  
QY 421 GGGGCTCCCTTATGGGGGCGTGCAGAGGCGCTGGGCGATGGCGTCCGGGTTGAGAGAC 480  
DB 421 GGGGCTCCCTTATGGGGGCGTGCAGAGGCGCTGGGCGATGGCGTCCGGGTTGAGAGAC 480  
QY 481 GGGGTGAATATGCAACAGGAATTTACCGGTTGCTTTCTATCTTCTCTTGGCT 540  
DB 481 GGGGTGAATATGCAACAGGAATTTACCGGTTGCTTTCTATCTTCTCTTGGCT 540

QY 541 TTGCTGCTCTGTTTGACCAATTCAGCTTCCGTTATGAAGTGCACAAGTGTCCGAGATC 600  
DB 541 CTGCTGCTCTGCTGACCAACCCAGCTTCCGTTATGAAGTGCACAAGTGTCCGAGATC 600  
QY 601 TACCATGTGACGAAGATGTTCTCAATCAGACATGTTGTACAGACAGCGGACATGATC 660  
DB 601 TACCATGTGACGAAGATGTTCTCAATCAGACATGTTGTATGAGGACAGCGGACATGATC 660  
QY 661 ATGCAACCCCCGGGTGTGCTGTGTGTCGGGAGGGTAAATTCCTCCCGCTGCTGGGTA 720  
DB 661 ATGCAATACCTCCGGGTGTGCTGTGTGTCGGGAGGGCAATCTCTCCGCTGCTGGGTA 720  
QY 721 GGGCTACTCCACGCTGCGGCGCAAGAGCGCAGATCCCGCATGCGCAATACAGACGC 780  
DB 721 GGGCTACTCCACGCTGCGGCGCAAGAGCGTACATCTCCACAGACGATACAGACGC 780  
QY 781 CAGCTGATTTGCTGTTGGGCGGCTGCTTCTGCTCGCTATGATGCTGGGGGATCTC 840  
DB 781 CAGCTGATCTGCTGTTGGGCGGCTGCTTCTGTTCCGCTATGATGCTGGGGGATCTC 840  
QY 841 TGGGATCTGTTTCTGCTGCTCTGAGCTGTTACCTTCTGCTGCGGATCAGAC 900  
DB 841 TGGGATCTGTTTCTGCTGCTCTGAGCTGTTACCTTCTGCTGCGGATCAGAC 900  
QY 901 GTACAGGACTGCAATTTGTCATCTATCCCGGCGAAGTATCAGGTACCGGATGGCTTGG 960  
DB 901 GTACAGGACTGCAATTTGTCATCTATCCCGGCGAAGTATCAGGTATCGGATGGCTTGG 960  
QY 961 GATATGATGATGAATGATGCTCACTTACAGAGCTTATGATGATTCGAGTACTCGGATC 1020  
DB 961 GATATGATGATGAATGATGCTCACTTACAGAGCTTATGATGATTCGAGTACTCGGATC 1020  
QY 1021 CCAACAAGCTGTGTGACATGTTGGGGGGGCGCCACTGGGGAGTCTTGGCGGCTTGGC 1080  
DB 1021 CCAACAAGCTGTGTGACATGTTGGGGGGGCGCCACTGGGGAGTCTTGGCGGCTTGGC 1080  
QY 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
DB 1081 TACTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
QY 1141 GTTGAAGGGAACCTTACAGACAGAGGGGGAACACAGCGCGCGCCACGCGGCTTACA 1200  
DB 1141 GTTGAAGGGAATACCAAGTGAACAGGGGGGGGGAACCAACCAACAGCGGCTCGTG 1200  
QY 1201 TCCCTCTTCACTGAGCGGCTGCAAAATTCAGCTTGTAAACCAACAGCGGCTCGTG 1260  
DB 1201 TCCATGTTGCAAGTGGGCGCTGCAAAATTCAGCTTGTAAACCAACAGCGGCTCGTG 1260  
QY 1261 CACATCAACAGAACTGCTTGAATGCAATGATCTCCCTCAAACTGGGTTCTTGGCGG 1320  
DB 1261 CACATCAACAGAACTGCTTGAATGCAATGATCTCTTCAACATGGGTTCTTGGCGG 1320  
QY 1321 CTGTTCTACACGACAGGTTCAATGCTGCTCGATGCTCAGAGCGGATGCGGCTGCGC 1380  
DB 1321 CTGTTCTACACAGATGTTCAATGCTGCTCGATGCTCAGAGCGGATGCGGCTGCGC 1380  
QY 1381 CCCATTGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440  
DB 1381 ACCATTGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440  
QY 1441 GACAGAGGCGCTATGCTGGGACATACGCACTCAACCGTGTGATGATGATGATGATG 1500  
DB 1441 GACAGAGGCGATATGCTGGGACATACCACTCAACCAATGATGATGATGATGATGATG 1500  
QY 1501 CAGGTGTGTGGCCAGTGTATGTTCACTCCAAAGCCCTGTGTGTGTGGGAGCAGCAGAT 1560  
DB 1501 CAGGTGTGTGGCCAGTGTATGTTCACTCCAAAGCCCTGTGTGTGTGGGAGCAGCAGAT 1560  
QY 1561 CGTTTGGCGCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620  
DB 1561 CGTTTGGCGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620

```
Qy 1621 AACACGCGGCCGCAACGAGGCACTGTTGGCTGTACATGATGATAGCACCGGGTTC 1680
    |||||
Db 1621 AACACGCGGCCGCGCGAAGGCACTGTTGGCTGTACATGATGATAGCACCGGGTTC 1680
    |||||
Qy 1681 ACCAAGACGTGTGGGGGCCCCCGTGCAACATCGGGGGGTTCGCAACAACCTTTGATC 1740
    |||||
Db 1681 ACCAAGACATGTGGGGGGCCCCCGTGTAACATCGGGGGGTTCGCAACAACACCTTGACC 1740
    |||||
Qy 1741 TGCCCCACGACTGCTTCGGAAGCATCCGAGGCCACTTACACCAAAATGCGTTGCGGG 1800
    |||||
Db 1741 TGCCCCACGACTGCTTCGGAAGCACCCCGAGGCTACTTACACCAAAATGTGTTGCGGG 1800
    |||||
Qy 1801 CCTTGGTTG 1809
    |||||
Db 1801 CCTTGGCTG 1809
    |||||
```

Search completed: February 19, 2005, 11:51:43  
Job time : 316.302 secs



C	1	137	7.6	488	1	AV7558731	AV7558731
C	2	104.4	5.8	492	1	AV7589366	AV7589366
C	3	58	3.2	179	7	R28798	R28798 F0-26d 22
C	4	56		844	9	CNS005012P	AL056652 Drosophil
C	5	55.4	3.1	925	9	CNS00931P	AL053013 Drosophil
C	6	51.8	2.9	879	9	CNS0200G	AL209545 Tetradom
C	7	50.2	2.8	1036	9	CNS010BS	AL088770 Drosophil
C	8	49.8	2.8	925	9	CNS0091P	AL053013 Drosophil
C	9	49.4	2.7	1101	9	CNS010UE	AL107216 Drosophil
C	10	46.4	2.6	1030	9	AG126333	AG126333 Pan trogl
C	11	46	2.5	1100	9	CNS015BR	AL106855 Drosophil
C	12	45.4	2.5	1201	9	CNS015BR	AL106455 Drosophil
C	13	45.2	2.5	839	9	CNS004NB	AL054280 Drosophil
C	14	44.2	2.4	910	9	CNS006ON	AL056529 Drosophil
C	15	44.2	2.4	1101	9	CNS006NG	AL057398 Drosophil
C	16	44	2.4	932	9	CNS0072Q	AL066742 Drosophil
C	17	43.6	2.4	997	9	CNS006DN	AL065132 Drosophil
C	18	43.6	2.4	1009	9	CNS010BW	AL098882 Drosophil
C	19	43.2	2.4	1058	5	BX462094	BX462094
C	20	42.8	2.4	1071	9	CNS006EY	AL069335 Drosophil
C	21	42.6	2.4	645	9	CNS01213	AL101589 Drosophil
C	22	42.6	2.4	932	9	CNS0072Q	AL066742 Drosophil
C	23	42.6	2.4	936	4	BG852371	BG852371 1024034A0
C	24	42.2	2.3	731	6	CB651667	CB651667 OSJNB1616

C	25	41.8	2.3	442	1	AF6367693
C	26	41.6	2.3	993	5	BX383247
C	27	41.6	2.3	1135	5	BW527635
C	28	41.4	2.3	864	9	CNS01219
C	29	41.4	2.3	439	7	CO141650
C	30	41.2	2.3	450	7	CO137934
C	31	41.2	2.3	743	7	CO138550
C	32	40.8	2.3	356	5	BY338946
C	33	40.8	2.3	628	8	BZ6363741
C	34	40.8	2.3	642	3	CR721023
C	35	40.8	2.3	774	8	BZ797506
C	36	40.8	2.3	937	8	BZ638670
C	37	40.8	2.3	855	9	CC632545
C	38	40.8	2.3	867	9	CC632352
C	39	40.8	2.3	914	9	CG977474
C	40	40.8	2.3	997	9	CG286439
C	41	40.6	2.2	359	4	BJ252669
C	42	40.6	2.2	375	4	BJ246716
C	43	40.6	2.2	518	9	CNS016DP
C	44	40.6	2.2	970	9	CNS010C9
C	45	40.4	2.2	409	5	BQ469708
						AF367693
						BX383247
						BW527635
						CNS01219
						CO141650
						CO137934
						CO138550
						BY338946
						BZ6363741
						CR721023
						BZ797506
						BZ638670
						CC632545
						CC632352
						CG977474
						CG286439
						BJ252669
						BJ246716
						CNS016DP
						CNS010C9
						BQ469708
						H2OIKOLF

## ALIGNMENTS

FEATURES	REFERENCE	ORGANISM	LOCUS	RESULT 1
source	Unpublished (2000)	Homo sapiens	AV755731	AV755731/c
location/Qualifiers	Contract: Zeguang Han	Homo sapiens (human)	AV755731.1	AV755731
1. 488	Chinese National Human Genome Center at Shanghai	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	GI:10913579	488 bp mRNA linear EST 19-OCT-2000
/organism="Homo sapiens"	351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai			
/mol_type="mRNA"	201203, P. R. China			
/cd_xref="taxon:9606"	Tel: 86-21-50801919(ex.45)			
/clone="EMFARKB03"	Fax: 86-21-50801922			
/rissue_type="Bone marrow"	Email: hanzg@chgc.sh.cn			
/cell_type="CD34+ hematopoietic stem/progenitor cell"	This clone is available at CHGC in Shanghai.			
/lab_host="BM25.8"				
/clone_1lb="BM"				
/note="Vector: pTiripEx2; Site_1: sfilA; Site_2: sfilB"				
ORIGIN				
Query Match				
Best Local Similarity	7.6%; Score 137; DB 1; Length 488;			
Matches	280; Conservative 0; Mismatches 182; Indels 5; Gaps 3;			
db	472 GTGGTGTGCACACTGCCTCTCTAGCTCTACCTGTGGGACCTCTTCGACGAGTATG			
796 GTTGGGCGGCTGCTTCTGCTCGGCTATGATGATGCGGGAGTCTTCGCGATCTGTTTC				
413				

QY 856 CTCGCTCTCAGACTGTTCACCTTCTGCGCTCGCCGACATCAGACCGGTACAGACTGCAT 915  
 DB 412 CTTCGACGTTTCAGCTGATCA--TCTGCTCTCAGACCACTGATTTGTGATGAAATGCAAC 356  
 QY 916 TGTTCATCTATCCCGGCAAGATATCAGATCAACCGCATG-GCTTGGATATGATGATGAA 974  
 DB 355 TGTCTCATCTATCTGCGGCGCATCTGGAACCGGTATGAGCATGGACATGATGATGAA 296  
 QY 975 CTGTGCACCTTACAGACCGCTAGTGTATTCGAGTACTCCGGATCCCAAGCTGTGTGT 1034  
 DB 295 CTGTGTGTCACCGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 236  
 QY 1035 GACATGTGTGCGCGGCGCCCACTGCGGAGATCTGCGGCGCTTGTCTTATTCATGAT 1094  
 DB 235 AGATATCATCAGCGGCGCTCACTGCGGCGCTATGTTGCGCTTACGCTTCTCTATGCA 176  
 QY 1095 GGGGAATCTGAGCTTAAAGTCTTGTGTGTGATGATGATGATGATGATGATGATGATGAT 1154  
 DB 175 GGGAGCGTGGCGGAAGTCTGCTGATCTTCTGTTGACCGCTGCGGATACAGCGAGCAC 116  
 QY 1155 TTAACGACAGCGGGGAC-AACGCGCGCGCGCGCCACGCGGTATTCATCCCTTTCACAC 1213  
 DB 115 CCACGCTCAGAGTGGCGCTGTAAGCGCGCGCGCGCGAGATTTGCTCAGCTTGTGCTACC 56  
 QY 1214 CTGGCGCGGCTCAGAAATCCAGCTTGTAAACCAACGCGAGCTGG 1260  
 DB 55 CTGGCGCTTAAACAGAACATCCACATCATTTANATGATTCGATGG 9

RESULT 2  
 LOCUS AV758366 492 bp mRNA linear EST 19-OCT-2000  
 AV758366/c  
 DEFINITION AV758366 BM Homo sapiens cDNA clone BMFKA03 5', mRNA sequence.  
 ACCESSION AV758366  
 VERSION AV758366.1 GI:10916214  
 KEYWORDS EST.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 492)  
 Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,  
 Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,  
 Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and  
 Chen,Z.  
 Homo sapiens cDNA BM clones  
 Unpublished (2000)  
 JOURNAL  
 COMMENT Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919 (ex. 45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 location/Qualifiers  
 1..492  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="BMFKA03"  
 /issue\_type="Bone marrow"  
 /cell\_type="CD34+ hematopoietic stem/progenitor cell"  
 /lab\_host="BM25.8"  
 /clone\_lib="BM"  
 /note="Vector: pTriplex2; Site\_1: sf1A; Site\_2: sf1B"

ORIGIN  
 Query Match 5.8%; Score 104.4; DB 1; Length 492;  
 Best Local Similarity 56.1%; Pred. No. 8.7e-17;  
 Matches 254; Conservative 0; Mismatches 194; Indels 5; Gaps 3;

QY 812 TCTGCTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 871  
 DB 457 TGTATCAGCTCATCTAGCTGTGTGACCTCTGCGTGTGGGATATGCTTGCAGCCCACTGA 398  
 QY 872 TCACCTTCTGCGCTGCGCATCAGACGATGACAGACTGATGATGATGATGATGATGATGATGAT 931  
 DB 397 TTA--TCTTCACAGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 341  
 QY 932 GCCACGATACAGTTC-ACCGATGCGTGGATGATGATGATGATGATGATGATGATGATGATGAT 990  
 DB 340 GTCATCATCAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 281  
 QY 991 GCCCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050  
 DB 280 TCCATGATACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 221  
 QY 1051 GCCCACTGGGAGATCTGCGCGGCTTGTCTTATTCATGATGATGATGATGATGATGATGAT 1110  
 DB 220 GCACACTGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 161  
 QY 1111 GCTTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1170  
 DB 160 GTACCTGTATCATCTGTTTACCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 101  
 QY 1171 ACACAGCGCGCGCGCGCCACAGCGGCT-TACATCCCTCTTACACCTGCGCGCGCTCAGAA 1229  
 DB 100 ATCACTGGCGCGCGCGCGAGACAGAGATTTGTGATGATGATGATGATGATGATGATGATGAT 41  
 QY 1230 AATCAAGCTTGTAAACCAACGAGCGAGCTGCA 1262  
 DB 40 CATCAGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8

RESULT 3  
 LOCUS R28798 179 bp mRNA linear EST 03-DEC-1999  
 R28798  
 DEFINITION F0-266D 22 week old human fetal liver cDNA library Homo sapiens  
 R28798  
 ACCESSION R28798  
 VERSION R28798.1 GI:6514170  
 KEYWORDS EST.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 179)  
 Choi,S.S., Yun,J.W., Choi,E.K., Cho,Y.G., Sung,Y.C. and Shin,H.S.  
 Construction of a gene expression profile of a human fetal liver by  
 single-pass cDNA sequencing  
 Mamm. Genome 6 (9), 653-657 (1995)  
 JOURNAL  
 COMMENT Contact: Hee-Sup Shin  
 Developmental Genetics  
 Pohang Institute of Science & Technology  
 San31, Hyodang Pohang, 790-784 Republic of Korea  
 Tel: 562-279-2291  
 Fax: 562-279-2199  
 Email: shinh@vision.postech.ac.kr  
 Seq primer: T3 primer.  
 location/Qualifiers  
 1..179  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="F0-266D"  
 /lab\_host="XLA-blue MRP"  
 /clone\_lib="22 week old human fetal liver cDNA library"  
 /note="Vector: pBluescriptII SK(-); Site\_1: EcoRI; Site\_2:  
 XhoI; The cDNA library made by oligo-dT primed and  
 directionally cloned between 5' Exor I-XhoI3' sites."

ORIGIN  
 Query Match 5.8%; Score 104.4; DB 1; Length 492;  
 Best Local Similarity 56.1%; Pred. No. 8.7e-17;  
 Matches 254; Conservative 0; Mismatches 194; Indels 5; Gaps 3;

Query Match	3.2%	Score 58;	DB 7;	Length 179;
Best Local Similarity	86.5%;	Pred. NO. 0.00024;		
Matches	64;	Conservative	0;	Mismatches 10;
				Indels

0y 1223 CTAGAAATCCAGCTTGTAAACACCGGCACTGGGACATCAACAGAACTGCTTGA 1282

Db 63 CGCGAACCTTCAAGTCTGTAAACGCTAATGGCACTGGGCATCAACAGACTGCTTGA 122

QY	1283	ACTGCATGACTCC	1296
Db	123	ATTGCAATGAGTCC	136

## RESULT 4

LOCUS	CNS0052P	844 bp	DNA	linear	CGS 03-JUN-1999
DEFINITION	Drosophaster melanogaster genome survey sequence TET3 end of BAC #				
	BACR116 of RPCT-98 library from Drosophila melanogaster (fruit fly); genomic survey sequence.				
ACCESSION	AF054652				

VERSION	AL056652.1	GI:4932342
KEYWORDS	GSS.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	

REFERENCE  
1 (bases 1 to 844)  
Eukaryota; Metazoa; Alveolipoda; Hexapoda; Insecta; Pterygota;  
Euphyptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephypteroidea; Drosophilidae; Drosophila.

**TITLE** Direct Submission  
**JOURNAL** Genoscope - Centre National de Séquençage :  
Submitted (02-JUN-1999) BP 191 91006 Evry cedex - FRANCE (E-mail : [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr)  
- web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))

**COMMENT**  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information

please see <http://www.itnity.org> The BAC Drosophila melanogaster BAC library was prepared by Kazutoko Oseegawa and Aaton Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

```

/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR1P16"
/clone_1fb="RPCI-98"
/notes="end : TET3"

```

## ORIGIN

Query Match	3.1%;	Score 56;	DB 9;	Length 844;
Best Local Similarity	21.6%;	Pred. No. 0.0012;		

Matches	76; Conservative	131; Mismatches	145; Indels	0; Gaps	0;
203	CTGCACGCCCGAGGCGACGGGCTGGGCTTCAGCCCGGGTACCCCTTGGCCCTTATGGCA	262			

Db 644 CCCCCCYSCCBBCBNCNTSCSYSYBSCSCSSSSSSSCSGSBSYSCCBCCBGCCG 585  
 Oy 263 ACGAGGGCATYGGGTGGGCGAGATGCTCTGTCACCCCGTTGGCTCCCGGCTTATTGG 322

Db 584 SBVCSBSBSSTBSVSGSBSSTSSGTCCKCBSSSSSBBBCSSBSGCCVCSSSYCTCYCSSC 525

Qy 323 GCCCAGTACCCCCGGCGGAGTCGCGTAATTTGGGTAAAGTATCGATCCGTCACAT 382

```
Db      BYCBSSGGBTBYSBSCSGBTBSSBSGGKSGGKSNGKGGBKSGGKGVICBYTCCSB8    465
Qy     GGCGCTTGCCACCTCAATGGGGGTACAATTCCGCTGCGAGCTCCTTTAGGGGGCGCTG   442
```

443 CCAGGGCCCTGGCGCATGGGCTCCGGGTTCTGAGAGACGGGCTGAATATGCAACAGGA 502

Db 404 YSTSSSSSGCGTGNKNTTGCVCSTCTTTKASCTBSAGSSBSSTATTGTTTTTTTTTG 345

Qy 503 ATTAAACCGGTGCTCTTTCTCTAATCTCTCTTTGCTTTGCTGCTGTTT 554

Db 344 TTTTTCCTTTTTTTTTTGTTT 293

## RESULT 5

LOCUS	CNS0091P	925 bp	DNA	linear	GEN 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
ACCION	BACRID16 of RPc1-58 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				

VERSION	AL053013.1	GI:4934461
KEYWORDS	GSS.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	

REFERENCE  
1 (pages 1 to 925)

MOTIONSCOPE  
 TITLE  
 JOURNAL  
 Direct Submission  
 Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))

**COMMENT**  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information

Please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Oosawa and Aaron Mammoss in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPC1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp. The same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

```
1.543 /organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"
```

**ORIGIN**

Query Match	3.1%;	Score 55.4;	DB 9;	Length 925;
Best Local Similarity	14.0%;	Pred No 0	0017.	

Matches	50; Conservative	172; Mismatches	134; Indels	1; Gaps
128	GGTTGGGTGGCGCGACATGAGAGACTTCGAGCGGTGCGCACTCGTGAAGCGAC	187		

Qy	Db
188	557
AACTATCCCCAAGCTCGGCAGCCCGAGGAGGAGGCTTGGGCTAGACCCGGTACCCTT	GGGKGCSSGSGSGBSCSCSSCSGSSSSGSCBCCCCCSBYCCSSBSBSKSSGTSBS
247	616

Db 617 CCCCCSKVCGTCGTCSSSSSCSSSSSSTSSSSTSSSTSKSSSSGSSSSSSSSSYTTSKSTS 676

Qy 248 GGGCCCTCTATGGCAACGAGGCATGGGGTGGGCGAGATGCTCCTGTCACCCCGTGGCT 3077

61 GACGTCAAGTTCCCGGGGGTGTGACATCGTTGGTGAAGTTTACCTGTTGCCGCGCAGG 120

```

Db      629 CCAACMAACACCCSSAVSMGSAANAAAAAASAAVAMGMAAAAAAGASCAGAGGR 608
Qy      121 GGGCCCAAGTGGTGTGCGCGGACTAGAGAATTCCAGCGGTCCCACTCTGTGA 180
Db      689 SSSSCCGGCGSVCCYCKCGCCGCCCGAAGAAAGACGAGSAV--GCGSAAAGCASSAGCA 745
Qy      181 AGCGCAACACTTCCCAAGGCTCGCCAGCCCGGAGGCGAGGCTTGGGCTCAGCCCGG 240
Db      746 MMGASCSAARVCCGCCGCGSCVSCGSCGSGMGSCGAMASSSSCSASMA 805
Qy      241 TACCTTGGCCCTCTATGCAACAGAGGATGGGATGGGATGGCTCTGTCAACC 300
Db      806 AGASAGSGGCGSCGSCASSVSSGSCGCMGCVSCAGACGCCSSCGSCGSSSGMGCSA 865
Qy      301 CTGTGCTCCCGGCTTATGAGGCGCCCACTGACCCCGCGGTAGTGGCTAAATTTGGT 360
Db      866 SGSGSASASSSGSTCGSSSSGCGASMAAGSGSGSVAAAGCGSADAGCRAG 925
Qy      361 AAGTATGATACCTTACATGCGGCTTGGCGGACTTATGGGTATCTCCGCTCTC 420
Db      926 SSGCGACNSCAGMCKCGMAAAASVSGMAAAAAAARASGCGSSSCSGSGSVGS 985
Qy      421 GCGGCTCCCTTAGGCGGCGCTGCGACAGGCGCTGCGC 457
Db      986 BBTCSARVARSGGCGSAGCGSSSGSSSVSGSS 1022

```

RESULT 8  
CNS0091P/c  
LOCUS  
DEFINITION  
925 bp DNA linear GSS 03-JUN-1999  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR19D16 of RCT-98 library from Drosophila melanogaster (fruit  
fly); genomic survey sequence.  
AL051013  
GI:4934461  
GSS.

ACCESION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster (fruit fly)  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 925)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

COMMENT  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osoegawa and  
Aaron Mammeter in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain y2; cn bw sp. the same strain used for the BDGP's  
p1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
source  
Location/Qualifiers  
1..925

1..925  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR19D16"  
/clone\_lib="RPI-98"  
/note="end : TET3"

ORIGIN

Query Match 2.8%; Score 49.8; DB 9; Length 925;  
Best Local Similarity 15.4%; Pred. No.0.056;  
Matches 65; Conservative 173; Mismatches 179; Indels 5; Gaps 1;

```

Qy      112 CCGCGAGAGGCGCCCAAGTTGGTGTGCGCGACTAGAGAATTCCAGCGGTGCA 171
Db      924 SBCSCSCSCSSBSCSSSSMSTSSNSBSCSSBSSSTSSMSSSSBSSSSGSSSS 865
Qy      172 CTTGTGGAAGGAGCAACTTATCCCAAGGCTTCCAGCCGAGGCGAGGCTGGACT 231
Db      864 SGTSAACVCKMASSSGCCGCGABCMCSSSSSCCGSARVXYRASGAKRGGSG 805
Qy      232 CAGCCCGGATACCTTGGCCCTCTATGCAACAGAGGATGGGTGGAGATGGCTC 291
Db      804 GASHSSSSSACSSSSSCASCSWASSSSSASSSRSRGGAGSGAGSSRSSSSA 745
Qy      292 CTGTACCCCGGCTTCCCGGCTTATGAGGCGCCCACTGACCCCGCGGTAGTGGCT 351
Db      744 SAGSVSASSSSSSCSSSVSCSVASMSBSSSSASASSSSSSASCSGCCCT 685
Qy      352 AATTGGATTAAGTATCGATACCTTCACATGCGGCTTGGCGGACTCATGGGCTAT 411.
Db      684 SWSGSCSTSAWSMASSSSSSSSSSSSSASSSASASSSSSSSSSSGSA---- 629
Qy      412 CGGCTGTGCGGCGCTCCCTTAGGCGGCGGCGCCAGGCGCTGCGGATGGCGGCT 471
Db      628 -CGSSMSGGGSSVASSGMSVSSSGGSGGCGGCGGSSSGSSSGSGSV 570
Qy      472 CTGAGAGCGGCGCTGATGATGCAACAGGAAATTTACCCGTTGCTTTCTATCTTC 531
Db      569 CSCSSGCMCRCSASSAAAAAASCVASCGMCKSKSGCTNNNNNTNTNTNTNTNN 510
Qy      532 CT 533
Db      509 NT 508

```

RESULT 9  
CNS016U6  
LOCUS  
DEFINITION  
1101 bp DNA linear GSS 26-JUL-1999  
Drosophila melanogaster genome survey sequence T7 end of BAC  
BACN16G16 of DrosBAC library from Drosophila melanogaster (fruit  
fly); genomic survey sequence.  
AL107216  
GI:5626210  
GSS.

ACCESION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster (fruit fly)  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Genoscope.  
Direct Submission  
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

COMMENT  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (BDGP) -  
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billand at CRPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
and project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

FEATURES  
source  
Location/Qualifiers  
1..1101

1..1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACN16G16"  
/clone\_lib="DrosBAC"  
/plasmid="pBelobAC11"



Best Local Similarity 22.6%; Pred. No. 0.62;  
Matches 119; Conservative 167; Mismatches 239; Indels 1; Gaps 1;

```
QY 49 CGCCGCCACAGACGCTCAAGTTCCTCCGCGCGGTGTGATGCTGTGATGTTTACCTG 108
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 556 CGCTGYCMGTGGTGGKKGKGTGGGCKGKGTGGGSGKGTGGTGGTGGGSGCGCTGG 615
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 109 TTGCGCGGACGGGCGCCAGTGTGGGTGGCGGCACTAGGAAGACTTCCGAGCGGTG 168
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 616 TGTYYTKSGKGYGCGCCBGGSGGCKTCCGSSGGGGYKSKCSGSCCGCGCGSGSG 675
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 169 CAACCTGTGGAAGAGACAACTATCCCAAGGCTCCGACGCCGACGAGGCGAGGCTG 228
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 676 CGGSSSSSSGCGKCGSKSGSCSBBSGGYCCBGGSSSKCKCSBGGGSGGCGCGG 735
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 229 GCTGACCCCGGAGTACCTTGGCCCTCTATGCAACAGAGGCGATGGGGTGGGCAAGATG 288
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 736 CGCTCCGCGKBGGGCGYCGCGSKCBCTGTCCTCCSCYCTCKSKSCGCGTGTGCK 795
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 289 CTCCTTACCCCGGTGCTCCGCGCTAGTTGGGCGC-CCCACTGACCCCGGCGTAGTC 347
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 796 CKCTGTGCKBYCKCTTCKYSYCKBCYKCGCGCGSKCSCKGCKCKTKCKY 855
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 348 GCGTAATTTGGGTAAGTCAATGATACCTCAATGCGGCTTGGCGGCACTCATGGGTA 407
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 856 KKKKKKKKKKKKKCKCTYKCKKCBCKCKYKCKBKCKCKBKCKCKCKCKCKCKBK 915
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 408 CATTCGCTGTGCGGCTCCCTTAAGGAGCGGTGCGAGGCGCTGCGGATGAGCGCTCG 467
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 916 BKCBKCBKCKCKCBCKBCCBKCKYKCKCBCKCKYKCKBKCKYKCKCBCKCK 975
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 468 GGTTCGAGAGAGCGGTGAATATGCAACAGGAATTTACCGGTGCTCTTCTAT 527
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 976 KCCCBCKCBKCKCKCKCKCKCKCKCKCKCKBKCBCKCBCKCBCKCBCKCKKYT 1035
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

RESULT 12
CNS016BR 1201 bp DNA linear GSS 26-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION BACN15K14 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL106545
VERSION AL106545.1 GI:5622456
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1201)
REFERENCE
Direct Submission
Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES
source
1..1201
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
```

## ORIGIN

Query Match 2.5%; Score 45.4; DB 9; Length 1201;

Best Local Similarity 18.4%; Pred. No. 0.92; Length 1201;

Matches 62; Conservative 135; Mismatches 140; Indels 0; Gaps 0;

/clone="BACN15K14"  
/clone\_11b="DrosBAC"  
/plasmid="pBelobAC11"  
/note="end : T7"

```
QY 36 ACCTAACACCAACCGCCGACAGAGCTCAAGTTCCTCCGCGGTGTGATGCTGTG 95
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 854 ANCTGMAAMMMMSGSGSCAAMVVVVVVVVVVVSVCCCMAGACHTCATKCCMG 913
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 96 TGAATTACCTGTTCTCCGCGAGGCGCCAGTGTGGTGGCGGCACTAGGAAGAC 155
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 914 DGASCCGDSVSSGGGCGCYRASSVGMHYATGGGKBCGCGATTWVSCSSGSSSCGKT 973
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 156 TTCCGAGCGGTGCAACTGTGGAAGGAGCAACTATCCCAAGGCTGCGGACGCCGA 215
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 974 VSGSSSSCCSSGSCCBCTGSSSSSSSCBCTSTCTYSSSSSSSKCSKSTSTY 1033
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 216 GGGCAGAGGCTGAGCTCAGCCGCGGATACCTTGGCCCTCTATGCAACAGAGGCA 275
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1034 SKGTCTSCYCKGTGTGKGGCCCGGSGCYCBGCCSKCKCGBSGTSKGTGKGS 1093
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 276 GTGGCAGAGATGCTCTCTACCCCGTGCCTCCGCGCTAGTTGGGCGCCCACTGACC 335
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1094 SSKGCGGCGSSSSSSSBTSBBSNTSSSTTTBBSSTSTTSBSGSHBTBBSBST 1153
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 336 CCGGCTAGTGCCTGTAATTTGGTAAATCAAT 372
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1154 STSBTBTSTTBTSSSBSTSTSSKKBGSSSTSSST 1190
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

RESULT 13  
CNS004NB/C 839 bp DNA linear GSS 03-JUN-1999  
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #  
DEFINITION BACR10E16 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION AL054280  
VERSION AL054280.1 GI:4931788  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly).  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see http://www.fruitfly.org The BDGP Drosophila  
melanogaster BAC library was prepared by Kanutoyo Osoegawa and  
Aaron Mammeser in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain Y2; cn bw ap, the same strain used for the BDGP's  
p1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.

COMMENT  
Direct Submission  
Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelobAC11.

FEATURES  
source  
1..839  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"







**THIS PAGE BLANK (USPTO)**